

... search, using sw model
... 31 32 seconds

Run on: January 10, 2002, 02:01:16 ; Search (without alignments)
OM protein 68.586 Million cell updates/sec

Run on:

US-09-712-768-8

File: 142
Perfect score: 1 DSIFLLVCLIRILRGAMSAKOHV
Sequence: 1

Scoring table: BLOSUM62, Gapext 0.5
Gapop 10.0

522463 seqs, 74073290 residues

Searches.
Total number of hits satisfying chosen parameters.

Minimum	DB seq	length:	0
	DB seq	length:	20000000000

Maximum DB seq	0%
Maximum Match	100%
post-processing:	
Maximum Match	45 summaries

Listing first 40

```
Database : A_Geneset-1
1: /SIDS2/gcdata
/SIDS2/gcdata
```

```
2: /SIDS2/500
3: /SIDS2/gcgcdata
3: /SIDS2/mcgcdata
```

```

:
A_Geneseq_1101.*
1: /SIDS2/gcgdata/geneseq/geneseqp/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseqp/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseqp/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseqp/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseqp/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseqp/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseqp/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseqp/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseqp/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseqp/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseqp/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseqp/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseqp/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseqp/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseqp/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseqp/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseqp/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseqp/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseqp/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseqp/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseqp/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseqp/AA2001.DAT.*

```

SUMMARIES

Result	Score	Query Match	Length	DB	ID	Description
NO.						
1	142	100.0	29	22	AAV97753	G. oxydans cytochr
2	66	46.5	133	21	AAAB29909	Human secreted pro
3	66	46.5	203	22	AAAB75603	Gene 46 human secr
4	65	45.8	62	21	AAAB34657	Human secreted pro
5	65	45.1	62	21	AAAB34658	Human secreted pro
6	64	45.1	78	22	AAAB75551	Human prostate can
7	64	45.1	93	22	AAAB63840	Human secreted pro
8	64	45.1	133	21	AAAB29910	Human secreted pro
9	64	45.1	203	22	AAAB75604	Human secreted car
10	61	43.0	35	21	AAAB57169	Human prostate can
11	58	40.8	93	22	AAAG75934	Human colon cancer

us-09-712-768-8.rag

13	46.5	32.7	94.5	180	22	AAW42392
12	46.5	32.4	94.5	20	AAV34594	
11	46	32.4	819	20	AAV34580	
10	46	32.4	3896	21	AAV33617	
9	46	32.4	3898	21	AAV53615	
8	46	32.4	3988	21	AAV53616	
7	46	32.0	1118	21	AAV58482	
6	45.5	32.0	115	21	AAV75789	
5	45.5	31.7	165	21	AAV7581	
4	45	31.7	183	21	AAV70291	
3	45	31.7	432	22	AAW40128	
2	44.5	31.3	503	22	AAW41914	
1	44.5	31.3	511	22	AAW08906	
	44.5	31.0	501	21	AAV72091	
	44	31.0	534	22	AAW74707	
	44	31.0	580	22	AAW93018	
	44	31.0	706	21	AAW06696	
	44	31.0	3898	12	AAW10473	
	44	31.0	3898	15	AAW05043	
	44	31.0	3898	17	AAW62529	
	44	31.0	3898	22	AAW20382	
	44	31.0	3898	22	AAW06462	
	44	31.0	352	20	AAW36224	
	43.5	30.6	353	20	AAW23023	
	43.5	30.6	378	21	AAW47435	
	43.5	30.6	378	21	AAW49751	
	43.5	30.6	378	22	AAW18247	
	43.5	30.3	127	21	AAW91277	
	43	30.3	278	22	AAW07067	
	43	30.3	308	21	AAW53865	
	43	30.3	308	21	AAW42866	
	43	30.3	325	21	AAW07075	
	43	30.3	361	21	AAW53866	
	43	30.3	361	21		

ALIGNMENTS

Human polypeptide
C. pneumoniae prot
porphyrinoma ging
porphyrinoma ging
Amino acid sequenc
NADP protein encond
Amino acid sequenc
Amino acid sequenc
Neisseria gonorrhoe
Neisseria meningit
Sequence encoded b
Human polypeptide
Human polypeptide
Human secreted prote
Human serine proteas
Human membrane ass
Human protein sequ
Human protein charact
protein characteri
Hog cholera virus
55 kilodalton prot
Hog cholera virus
Hog cholera virus
Hog cholera virus
Human Brainlaid pro
Human secreted prote
Human PRO424, plac
The beta-1,5-galact
Human beta-1,5-N-A
Arabidopsis thaliana
C glutamicum prote
Arabidopsis thaliana
Arabidopsis thaliana
Rhesus monkey meila
Arabidopsis thaliana
Arabidopsis thaliana

Page 1

RESULT	1
AAV97753	AAV97753 standard; Protein: 29 AA.
XX	
XX	
XX	
AC	AAV97753:
XX	06-AUG-2001 (first entry)
DT	
XX	G. oxydans cytochrome C oxidase (COII) protein sequence fragment.
DE	
XX	Cytochrome C oxidase; COI; COII; COIII; COI complex; D-sorbitol;
KW	oxidative fermentation; electron transfer; respiratory chain; L-sorbose
KW	2-keto-L-gluconic acid; 2KG production; aldehyde production.
KW	carboxylic acid production; ketone production.
XX	
XX	Gluconobacter oxydans.
OS	
XX	EP1103603-AZ.
PN	
XX	30-MAY-2001.
PD	
XX	14-NOV-2000; 2000EP-0124785.
XX	
PF	14-NOV-2000; 2000EP-0122842.
XX	
XX	17-NOV-1999; 99EP-0122842.
PR	
XX	(HOFF) HOFFMANN LA ROCHE & CO AG F.
XX	
PA	(HOFF) HOFFMANN LA ROCHE & CO AG F.
XX	
XX	Hoshino T, Shinoj M;
XX	
PT	Asakura A,
XX	
XX	WPI: 2001-357953/38.
DR	N-PSDB: AAA91492.
XX	New cytochrome C oxidase complex having cytochrome C oxidase activit
XX	from Gluconobacter oxydans DSM 4025, useful in mediating electron
PT	

PT transfer in respiratory chain or producing 2-keto-L-gulononic acid from
 PT L-sorbose or D-sorbitol
 PS Claim 6; Page 26; 42pp; English.
 CC This sequence represents a fragment of the gluconohexaric acid from
 CC cytochrome C oxidase, COXI. The invention relates to a cytochrome C
 CC ferredoxin complex. The COI complex is useful in improving oxidative
 CC in the respiratory chain. The recombinant mediating electron transfer
 CC (2KGA) from L-sorbose or D-sorbitol in a culture preparation of a
 CC acceptor from an enzyme belonging to dehydrogenase C, an electron
 CC especially the production of 2KGA from L-sorbose or D-sorbitol.
 SO Sequence 29 AA;

Query Match
 Best Local Similarity 100.0%; Score 142; DB 22; Length 29;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DSIFLVCILIRIRGAMSAKHVGFEMAA 29
 DB 1 DSIFLVCILIRIRGAMSAKHVGFEMAA 29

RESULT 2
 AAB29909

ID AAB29909 standard; Protein: 133 AA.
 AC AAB29909;
 DT 09-FEB-2001 (first entry)

Human secreted protein BLAST search protein SEQ ID NO: 167.
 KW Cytostatic; immunosuppressive; neutrotropic; neuroprotective; antiviral;
 KW anti-allergic; hepatotropic; antidiabetic; antidiarrheal; antitumor;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 OS Homo sapiens.
 PN WO200061779-A1.
 PD 19-OCT-2000.
 PF 06-APR-2000; 2000MO-US09068.
 PR 09-APR-1999; 99US-0128659.
 PA 20-JAN-2000; 2000US-0177050.
 PI (HUMA-) HUMAN GENOME SCI INC.
 DR Rosen CA, Ruben SM, Komatsoulis G;
 WP: 2000-647424/62.
 PT Isolated nucleic acid molecule encoding a human secreted protein is
 PT used in preventing, treating or ameliorating a medical condition -
 PS Disclosure; Page 487; 495pp; English.

The invention relates to the isolation of genes AAC63410-C63458 encoding
 CC peptide fragment homologous to the protein encoded by the gene given
 CC in the descriptor line. The sequence is a search result from a BLASTX
 CC biology search. The genes and proteins are useful for preventing,

CC ameliorating or treating medical conditions, e.g. by protein or gene
 CC therapy. The genes are isolated from a range of human tissues disclosed
 CC in the specification. The nucleic acids, proteins, antibodies and
 CC (ant)agonists are useful in the diagnosis, treatment and prevention of:
 CC adenoviral gland, bone, bone marrow, breast, gastrointestinal tract, liver,
 CC lung or urogenital; (b) immune disorders, e.g. Addison's disease,
 CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
 CC arthritis and ulcerative colitis; (c) cardiovascular disorders such
 CC as myocardial ischemias; (d) wound healing; (e) neurological diseases
 CC as viral, bacterial, fungal and parasitic infections.
 SO Sequence 133 AA;

Query Match
 Best Local Similarity 46.5%; Score 66; DB 21; Length 133;
 Matches 14; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

OY 2 SFIFLVCILIRIRGAMSAKHVGFEMAA 29
 DB 84 SFIFLVCILIRIRGAMSAKHVGFEMAA 111

RESULT 3
 AAB75603

ID AAB75603 standard; Protein: 203 AA.
 AC AAB75603;
 DT 06-APR-2001 (first entry)

Gene 46 human secreted protein homologous amino acid sequence #157.
 KW Human; secreted protein; immunosuppressive; antiarthritic; antitumor;
 KW antiproliferative; cytosolic; cardiant; vasotropic; cerebroprotective;
 KW neurotropic; neuroprotective; antibacterial; vituicide; fungicide;
 KW hyperproliferative; vulnerrary; autoimmune disease; cardiovascular disorder;
 KW nervous system disorder; ocular disorder; skin ageing; chemotaxis;
 KW food additive.
 OS Homo sapiens.
 PN WO200077026-A1.
 PD 21-DEC-2000.
 PF 01-JUN-2000; 2000MO-US14973.
 PR 11-JUN-1999; 99US-0138630.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI (ROSE/) ROSEN C A.
 DR Rosen CA, Ruben SM, Komatsoulis GA;
 WP: 2001-071258/08.
 PT Nucleic acid molecules encoding human secreted proteins, used in
 PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
 PT Parkinson's diseases and cancers -
 PS Disclosure; Page 94; 542pp; English.

Human secreted proteins AAB75506 - AAB75554 are encoded by polynucleotide
 CC sequences AAF64176 - AAF64224. The specification includes amino acid
 CC sequences AAB75555 - AAB75606 which represent fragments of the human
 CC secreted proteins, and protein sequences with which they share homology.
 CC The proteins and polynucleotides, their agonists and antagonists have
 CC activities dependent on the tissues and cells in which they are

PS Disclosure; page 405; 419pp; English

XX F
1000

XX
PI
Rosen CA, Ruben SM, Komatsoulis G;
XX
XX
XX
XX
XX

DR WPI: 2000-579482/54.

XX Isolated nucleic acid molecule encoding a human secreted protein is
PT used in preventing, treating or ameliorating a medical condition -
XX Disclosure: Page 405; 419pp; English.

CC The polynucleotide sequences given in AAC59738 to AAC59787 encode the
CC human secreted proteins given in AAB34577 to AAB34626. AAB34627 to
CC AAB34686 represent human secreted polypeptide sequences and proteins
CC homologous to them, which are given in the exemplification of the present
CC invention. Human secreted proteins have activities based on the tissues
CC and cells the genes are expressed in. Example of activities include:
CC antirheumatic; immunosuppressive; antineoplastic; antiproliferative;
CC cytostatic; cardiac; vasotropic; cerebroprotective; neurotrophic;
CC neuroprotective; antibacterial; virucide; fungicide; and
CC ophthalmological. The polynucleotides and proteins can be used to
CC prevent, treat or ameliorate a medical condition in e.g. humans, mice,
CC rabbits, goats, horses, cats, dogs, chickens or sheep. They are also
CC used in diagnosing a pathological condition or susceptibility to a
CC pathological condition. Disorders which are diagnosed or treated include
CC autoimmune diseases, hyperproliferative disorders e.g. neoplasms and
CC cerebrovascular disorders, angiogenesis, nervous system disorders,
CC infections caused by bacteria, viruses and fungi and ocular disorders.
CC The proteins can also be used to aid wound healing and epithelial cell
CC proliferation, to prevent skin aging due to sunburn, to maintain organs
CC before transplantation, for supporting cell culture of primary tissues
CC as a food additive or preservative in chemotaxis. The proteins can also be used
CC to regenerate tissues and in chemotaxis. The proteins can also be used
CC as a food additive or preservative to increase or decrease storage
CC capabilities. AAC59729 to AAC59737 and AAB34576 represent sequences used
CC in the exemplification of the present invention.

SQ Sequence 62 AA:

Query Match 45.1%; Score 64; DB 21; Length 62;
Best Local Similarity 50.0%; Pred. No. 0.0061;
Matches 14; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

OY 2 STELLVCLIRIRGMSAKOHGFEMAA 29
DB 13 stflfcifrlqllfhfshkfhgfeaa 40

RESULT 6

ID AAB75551 standard; Protein: 78 AA.
XX
AC AAB75551;

DT 06-APR-2001 (first entry)
XX

DE Human secreted protein sequence encoded by gene 46 SEQ ID NO:105.
XX

KW Human; secreted protein; immunosuppressive; antirheumatic;
KW antiproliferative; cytostatic; cardiac; vasotropic; cerebroprotective;
KW neurotrophic; neuroprotective; antibacterial; virucide; fungicide;
KW ophthalmological; vulnerrary; autoimmune disease; cardiovascular disorder;
KW hyperproliferative disorders; cerebrovascular disorder;
KW nervous system disorder; ocular disorder; skin ageing; chemotaxis;
KW food additive.

OS Homo sapiens.
XX

PN WO200077026-A1.
XX

PD 21-DEC-2000.
XX

PF 01-JUN-2000; 2000WO-US14973.
XX

PR 11-JUN-1999; 99US-0138630.
XX

PA (HUMA-) HUMAN GENOME SCI INC.
PA (ROSE/) ROSEN C A.
XX

PI Rosen CA, Ruben SM, Komatsoulis GA;
XX

DR WPI: 2001-071258/08.
DR N-PSDB: AAF64221.

PT Nucleic acid molecules encoding human secreted proteins, used in
PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
PT Parkinson's diseases and cancers -
XX

PS Claim 11; Page 490; 542pp; English.

CC Human secreted proteins AAB75506 - AAB75554 are encoded by polynucleotide
CC sequences AAF64176 - AAF64224. The specification includes amino acid
CC sequences AAB75555 - AAB75606 which represent fragments of the human
CC secreted proteins, and protein sequences with which they share homology.
CC The proteins and polynucleotides, their agonists and antagonists have
CC activities dependent on the tissues and cells in which they are
CC expressed, examples of these activities include: immunosuppressive;
CC antirheumatic; antineoplastic; antiproliferative; cytostatic; cardiac;
CC vasotropic; cerebroprotective; neurotrophic; neuroprotective; antibacterial;
CC virucide; fungicide; ophthalmological; and vulnerrary. The proteins,
CC polynucleotides, agonists and antagonists can be used to treat or detect
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders
CC e.g. neoplasms of the breast or liver, cardiovascular disorders
CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
CC angiogenesis, nervous system disorders e.g. Alzheimer's disease,
CC infections caused by bacteria, viruses and fungi and ocular disorders,
CC e.g. corneal infection. The polypeptides can also be used to aid wound
CC healing and epithelial cell proliferation, to prevent skin ageing due to
CC sunburn, to maintain organs before transplantation, for supporting cell
CC culture of primary tissues, to regenerate tissues and in chemotaxis. The
CC polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities. Included in the invention are
CC polynucleotide sequences AAF64167 - AAF64175 and peptide AAB75505 which
CC are used in the isolation, identification and characterisation of the
CC proteins of the invention.

SQ Sequence 78 AA:

Query Match 45.1%; Score 64; DB 22; Length 78;
Best Local Similarity 50.0%; Pred. No. 0.0078;
Matches 14; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

OY 2 STELLVCLIRIRGMSAKOHGFEMAA 29
DB 50 stflfcifrlqllfhfshkfhgfeaa 77

RESULT 7

ID AAB63840 standard; Protein: 93 AA.
XX
AC AAB63840;

DT 26-MAR-2001 (first entry)
XX

DE Human prostate cancer associated antigen protein sequence SEQ ID NO:1202.
XX

KW Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
KW cancer associated antigen; cytostatic; cancer vaccine.
XX

OS Homo sapiens.
XX

PN WO200073801-A2.
XX

PD 07-DEC-2000.
XX

PF 26-MAY-2000; 2000WO-US14749.
XX

PT Nucleic acid molecules encoding human secreted proteins, used in
PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
PT Parkinson's diseases -
XX
XX
Disclosure: Page 94; 542pp: English.
PS
TX

Human secreted proteins AAB75506 - AAB75554 are encoded by polynucleotide sequences AAF64176 - AAF64224. The specification includes amino acid sequences AAB75555 - AAB75606 which represent fragments of the human secreted proteins, and protein sequences with which they share homology. The proteins and polynucleotides, their agonists and antagonists have activities dependent on the tissues and cells in which they are expressed, examples of these activities include: immunosuppressive; antithalritic; antirheumatic; antiproliferative; cytostatic; cardant; vasotropic; cerebroprotective; nootropic; neuroprotective; antibacterial; vitucide; fungicide; opthalmological; and vulnerary. The proteins, polynucleotides, agonists and antagonists can be used to treat or diagnose various diseases and disorders including, autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiodysplasia, nervous system disorders e.g. Alzheimer's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection. The polynucleotides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin ageing due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities. Included in the invention are polynucleotide sequences AAF64167 - AAF64175 and peptide AAB75503 which are used in the isolation, identification and characterisation of the proteins of the invention.

Query Match	45.1%	Score 64	DB 22	Length 203
Best Local Similarity	50.0%	Pred. No. 0.022		
Matches 14	Conservative 3	Mismatches 11	Indels 0	Gaps 0
Qy	2	SIFLWCLRIIRLGANSARQHVGFEMAA	29	
Db	154	stfltcfcifqrlfnfscskhhgltcaaa	181	

RESULT	10	
AAB57169		
ID	AAB57169	standard; Protein; 35 AA.
XX		
AC	AAB57169;	
XX		
DT	13-MAR-2001	(first entry)
XX		
DE	Human prostate cancer antigen protein sequence SEQ ID NO:1747.	
XX		
KW	Human; prostate cancer; prostate cancer antigen; detection; diagnosis;	
KW	neuroprotective; cytosolic; cardioprotective; immunomodulatory; muscular;	
KW	vulnerable; gastrointestinal; nephrotropic; antineoplastic; gynaecological;	
KW	antibacterial; gene therapy; neural; immune; reproductive; renal;	
KW	gastrointestinal; pulmonary; cardiovascular; proliferative disorder;	
KW	wound; infectious disease.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200055174-A1.	
XX		
PD	21-SEP-2000.	
XX		
PF	08-MAR-2000; 2000WO-US05988.	
XX		
PR	12-MAR-1999; 99US-0124270.	
XX		
DA	(HUKA-) HUMAN GENOME SCI INC.	

PA (ROSE/ROSEN C A.
XX
XX Rosen CA, Ruben SM,
PI
XX WPI: 2000-587513/55.
DR N-PSDB; AAF16372.
XX
XX
PT Prostate cancer associated gene sequences, referred to as prostate
PT cancer antigens, useful for treatment, prevention, and diagnosis of
PT disorders such as prostate cancer -
XX
XX
PS Claim 11, Page 2216; 2338pp; English.
XX

CC AAF1556 to AAF16505 encode the human prostate cancer associated
CC proteins, called prostate cancer antigens, given in AAB56163 to AAB57302
CC The prostate cancer antigens can have neuroprotective, cytosstatic,
CC cardiolactic, immunomodulatory, muscular, vulnery, gastrointestinal,
CC nephrotropic, antinefctive, gynaecological and antibacterial activities
CC and can be used in gene therapy. The prostate cancer antigen
CC polynucleotides may be used for detection of prostate cancer.
CC Identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The prostate cancer antigens may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
CC AAF57303 represent sequences used in the exemplification of the present
CC invention.
XX
XX Sequence 35 AA;

```

Query Match      43.0%; Score 61; DB 21; Length 35;
Best Local Similarity 50.0%; Pred. No. 0.0097;
Matches 13; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 4 FLVCLIRILRCAMSKAKOHVFEMAA 29
    ||::||::||::||::||::||
Db 10 fltcltrqlrlfntskhnfgfeaaa 35

```

RESULT	11
AAG75934	
ID	AAG75934 standard; Protein; 93 AA.
AC	
AA	AAG75934;
DT	
XX	
DE	Human colon cancer antigen protein SEQ ID NO:6698.
XX	
KW	Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW	colorectal carcinoma; chromosome 17.
XX	
OS	Homo sapiens.
XX	
PN	WO200122920-A2.
XX	
PD	05-APR-2001.
XX	
PF	28-SEP-2000; 2000MO-US26524.
XX	
PR	29-SEP-1999; 99US-0157137.
PR	03-NOV-1999; 99US-0163280.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC.
XX	
PI	Ruben SM, Barash SC, Blrse CE, Rosen CA;
PT	WPI; 2001-235357/24.
DR	N-PDB; AAH35339.
XX	
PT	Nucleic acids encoding 4277 human colon cancer-associated polypeptides useful for preventing, diagnosing and/or treating colorectal cancers -

XX Claim 11; Page 8164; 9803pp; English.
 PS
 XX
 CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where
 CC the proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytostatic activity and can be used in gene
 CC therapy and vaccine production. N and P may be used in the prevention,
 CC diagnosis and treatment of diseases associated with inappropriate P
 CC expression. For example, N and P may be used to treat disorders
 CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of P by expressing
 CC inactive proteins or to supplement the patients own production of P.
 CC Additionally, N may be used to produce the colon cancer-associated P,
 CC by inserting the nucleic acids into a host cell and culturing the cell
 CC to express the proteins. N and P can be used in the prevention, diagnosis
 CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
 CC and AAB77789 represent sequences used in the exemplification of the
 CC present invention.
 CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were
 CC missing at time of publication, meaning no sequences are present for
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.
 XX
 SO Sequence 93 AA:

Query Match 40.8%; Score 58; DB 22; Length 93;
 Best Local Similarity 46.4%; Pred. No. 0.085; Mismatches 12; Indels 0; Gaps 0;
 Matches 13; Conservative 3;

OY 2 SIFLVLCIRILRGAMSAKOHVFEMAA 29
 DB 66 sftlfcirqlrlfhtskhfgfxtaa 93

RESULT 12

ID AAM42392 standard; Protein; 180 AA.

AC AAM42392;

DT 22-Oct-2001 (first entry)

DE Human polypeptide SEQ ID NO 125.

XX Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
 KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerrary;
 KW antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
 KW antineumatic; hepatotropic; cerebroprotective; antiinflammatory;
 KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
 KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; nephrotropic; gene therapy; vaccine.

OS Homo sapiens.

PN WO200155449-A1.

PD 02-AUG-2001.

PF 17-JAN-2001; 2001MO-US01346.

XX 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUL-2000; 2000US-0216880.
 PR 14-JUL-2000; 2000US-0218290.
 PR 14-AUG-2000; 2000US-0225447.
 PR 01-SEP-2000; 2000US-0229343.
 PR 06-SEP-2000; 2000US-0230437.
 PR 08-SEP-2000; 2000US-0231243.
 PR 25-SEP-2000; 2000US-0234997.
 PR 29-SEP-2000; 2000US-0236367.
 PR 13-OCT-2000; 2000US-0239937.

PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246528.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Barash SC, Ruben SM;

PI WPI: 2001-476225/51.

DR N-PDB; AAI62797.

XX Novel plasma membrane associated proteins useful for diagnosing,
 PT treating, preventing and/or prognosing disorders related to the
 PT proteins, including cancer, immune response and neuronal disorders
 XX
 PS Claim 11; SEQ ID NO 125; 532pp + Sequence Listing; English.

XX The invention relates to novel genes (AAI62752-AAI62961) and proteins
 CC (AAM42347-AAM42415) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
 CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
 CC and parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 180 AA:

Query Match 34.5%; Score 49; DB 22; Length 180;
 Best Local Similarity 27.3%; Pred. No. 4.7;
 Matches 6; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

OY 2 SIFLVLCIRILRGAMSAKOHV 23
 DB 127 avflvcvvnfkqrganehl 148

RESULT 13

ID AAY34694 standard; Protein; 945 AA.

AC AAY34694;

DT 13-SEP-1999 (first entry)

XX C. pneumoniae protein involved in metabolism of polypeptides.
 XX
 DE

KW Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
 KW sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
 KW vaccine; neutralising epitope.
 XX Chlamydia pneumoniae.
 OS
 PN WO927105-A2.
 XX
 PD 03-JUN-1999.
 PF 20-NOV-1998; 98WO-1B01890.
 PR 04-NOV-1998; 98US-0107078.
 PR 21-NOV-1997; 97FR-0014673.
 XX (GEST) GENSET.
 PA
 XX Griffais R;
 PI
 DR WPI: 1999-357842/30.
 XX
 PT Genome sequence of Chlamydia pneumoniae
 XX
 PS Page 692-694; Disclosure: 1912pp; English.
 CC AAY34584-Y35879 represent the proteins encoded by all the open reading
 CC frames in the complete genome (see AAX91990) of Chlamydia pneumoniae.
 CC C. pneumoniae causes respiratory disease such as pneumonia and
 CC bronchitis and is thought to be a contributing factor in heart
 CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema
 CC nodosum or pharyngitis. The polypeptides encoded by the open reading
 CC frames of the C. pneumoniae genome (see AAY34584-Y35879) can be used in
 CC immunogenic compositions as vaccines. Vectors containing C. pneumoniae
 CC nucleotide sequences can also be used as immunogenic compositions,
 CC especially where the vector directs the expression of a neutralising
 CC epitope of C. pneumoniae.
 CC
 XX
 SQ Sequence 945 AA;

Query Match 32.7%; Score 46.5; DB 20; Length 945;
 Best Local Similarity 42.9%; Pred. No. 73;
 Matches 9; Conservative 7; Mismatches 4; Indels 1; Gaps 1;

OY 3 IFLLV-CLIRILRGMSAKQH 22
 :|:| |:::| ||||
 Db 855 lfvlgyctalllgaaakqh 875

RESULT 14
 AAY34530
 ID AAY34530 standard; Protein; 750 AA.
 XX
 AC AAY34530;
 XX
 DT 25-AUG-1999 (first entry)
 XX
 DE Porphyromonas gingivalis protein PG61.
 XX
 KW Porphyromonas gingivalis; PG; periodontal disease; gingivitis;
 KW vaccine; antigenic.
 XX
 OS Porphyromonas gingivalis.
 XX
 PN WO929870-A1.
 XX
 PD 17-JUN-1999.
 PF 10-DEC-1998; 98WO-AU01023.
 PR 04-AUG-1998; 98AU-0005028.
 PR 10-DEC-1997; 97AU-0000839.
 PR 31-DEC-1997; 97AU-0001182.

PR 30-JAN-1998; 98AU-0001546.
 PR 10-MAR-1998; 98AU-0002264.
 PR 09-APR-1998; 98AU-0002911.
 PR 23-APR-1998; 98AU-0003128.
 PR 03-MAY-1998; 98AU-0003338.
 PR 22-MAY-1998; 98AU-0003654.
 PR 29-JUL-1998; 98AU-0004917.
 XX
 PA (CSLC-) CSL LTD.
 XX
 PI Agius CT, Barr IG, Hocking DM, Margetts MB, Patterson MA;
 PI Ross BC, Rothel LJ, Webb EA;
 XX
 DR WPI: 1999-385613/32.
 DR N-PSDB: AAX91748.
 XX
 PT Antigenic Porphyromonas gingivalis peptides for preventing
 PT gingivitis
 XX
 PS Claim 1; Page 520-521; 588pp; English.
 CC
 CC AAX91536 to AAX91801 encode two hundred and sixty six antigenic
 CC Porphyromonas gingivalis (PG) polypeptide sequences given in AAY34318 to
 CC AAY34583. AAX91802 to AAX91989 represent PCR primers used in the
 CC isolation of the PG polypeptides. The PG polypeptides have antibacterial
 CC activity with a vaccine mechanism of action. The PG polypeptides can be
 CC used as vaccines especially against Porphyromonas gingivalis. Probes can
 CC be used to detect Porphyromonas gingivalis in standard hybridisation
 CC assays. Porphyromonas gingivalis is involved in periodontal disease
 CC especially gingivitis.
 CC
 XX
 SQ Sequence 750 AA;

Query Match 32.4%; Score 46; DB 20; Length 750;
 Best Local Similarity 47.1%; Pred. No. 68;
 Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 12 ILRGMSAKQHVGEEMA 28
 ::|:| | ||||
 Db 635 vlagemlclshagfema 651

RESULT 15
 AAY34397
 ID AAY34397 standard; Protein; 819 AA.
 XX
 AC AAY34397;
 XX
 DT 25-AUG-1999 (first entry)
 XX
 DE Porphyromonas gingivalis protein PG61.
 XX
 KW Porphyromonas gingivalis; PG; periodontal disease; gingivitis;
 KW vaccine; antigenic.
 XX
 OS Porphyromonas gingivalis.
 XX
 PN WO929870-A1.
 XX
 PD 17-JUN-1999.
 PF 10-DEC-1998; 98WO-AU01023.
 PR 04-AUG-1998; 98AU-0005028.
 PR 10-DEC-1997; 97AU-0000839.
 PR 31-DEC-1997; 97AU-0001182.
 PR 30-JAN-1998; 98AU-0001546.
 PR 10-MAR-1998; 98AU-0002264.
 PR 09-APR-1998; 98AU-0002911.
 PR 23-APR-1998; 98AU-0003128.
 PR 05-MAY-1998; 98AU-0003338.
 PR 22-MAY-1998; 98AU-0003654.
 PR 29-JUL-1998; 98AU-0004917.

PR 29-JUL-1998; 98AU-0004917.

XX (CSLC-) CSL LTD.

XX Agius CT, Barr IG, Hocking DM, Margelits MB, Patterson MA;
PI Ross BC, Rothel LJ, Webb EA;

XX WPI; 1999-385613/32.
DR N-PSDB; AAX91615.

XX Antigenic Porphyromonas gingivalis peptides for preventing
PT gingivitis

XX Claim 1; Page 367-369; 588pp; English.

XX AAX91536 to AAX91801 encode two hundred and sixty six antigenic
CC Porphyromonas gingivalis (PG) polypeptide sequences given in AAY34318 to
CC AAY34583. AAX91802 to AAX91989 represent PCR primers used in the
CC isolation of the PG polypeptides. The PG polypeptides have antibacterial
CC activity with a vaccine mechanism of action. The PG polypeptides can be
CC used as vaccines especially against Porphyromonas gingivalis. Probes can
CC be used to detect Porphyromonas gingivalis in standard hybridisation
CC assays. Porphyromonas gingivalis is involved in periodontal disease
CC especially gingivitis.

XX Sequence 819 AA;

Query Match 32.4%; Score 46; DB 20; Length 819;
Best Local Similarity 47.1%; Pred. NO. 75;

Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 12 ILRGAMSAKOHVGFEMA 28
::|:| | |||||
Db 704 vlagsmllshagfema 720

Search completed: January 10, 2002, 02:01:17
Job time: 290 sec

Db 212 STFLTICFIRQLMFTSKHHFGEAAA 239

RESULT 2

; Sequence 2, Application US/08750717

; Patent No. 6180109

; GENERAL INFORMATION:

; APPLICANT: MOORMANN, Robertus J. M.

; APPLICANT: VAN RIJN, Petrus A.

; TITLE OF INVENTION: Nucleotide Sequences of Pestivirus

; TITLE OF INVENTION: Strains, Polypeptides Encoded by These Sequences and Use

; TITLE OF INVENTION: Thereof for Diagnosis and Prevention of Pestivirus

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: YOUNG & THOMPSON

; STREET: 745 South 23rd Street

; CITY: Arlington

; STATE: Virginia

; COUNTRY: USA

; ZIP: 22202

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/750,717

; FILING DATE: 24-DEC-1996

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: EP 94201743.5

; FILING DATE: 17-JUN-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: NO PCT/NL95/00214

; FILING DATE: 16-JUN-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: PATCH, Andrew J.

; REGISTRATION NUMBER: 32,925

; REFERENCE/DOCKET NUMBER: BO 39123

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 703-521-2297

; TELEFAX: 703-685-0573

; TELEX: 248425 EMBON

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3898 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-750-717-2

Query Match 33.1%; Score 47; DB 4; Length 3898;

Best Local Similarity 57.1%; Pred. No. 91;

Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 SIFLVCLIRLNG 15

Db 659 TIAFLICLIRKVRG 672

RESULT 3

; US-08-876-991-2

; Sequence 2, Application US/08876991

; Patent No. 5925360

; GENERAL INFORMATION:

; APPLICANT: Gregor Meyers, Tillmann R menapf,

; APPLICANT: Heinz-J rgen Thiel

; TITLE OF INVENTION: Hog cholera virus vaccine and diagnostic

; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Organon Teknika Corporation

; ADDRESSEE: Biotechnology Research Institute

; STREET: 1330-A Piccard Drive

; CITY: Rockville

; STATE: Maryland

; COUNTRY: U.S.A.

; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/876,991

; FILING DATE: 16-JUN-1997

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/747,577

; FILING DATE: US/08/650,584

; APPLICATION NUMBER: US/08/469,702

; FILING DATE: US/08/123,596

; APPLICATION NUMBER: 07/797,554

; FILING DATE: 22-NOV-1991

; APPLICATION NUMBER: US 07/494,991

; FILING DATE: 16-MAR-1990

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: William M. Blackstone

; REGISTRATION NUMBER: 29,772

; REFERENCE/DOCKET NUMBER:

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (301) 258-5200

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3898 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-876-991-2

Query Match 31.0%; Score 44; DB 2; Length 3898;

Best Local Similarity 70.0%; Pred. No. 2.8e+02;

Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 6 LVCLIRLNG 15

Db 663 LACLIKVRG 672

RESULT 4

; US-09-059-853-2

; Sequence 2, Application US/09059853

; Patent No. 5935582

; GENERAL INFORMATION:

; APPLICANT: Gregor Meyers, Tillmann R menapf,

; APPLICANT: Heinz-J rgen Thiel

; TITLE OF INVENTION: Hog cholera virus vaccine and diagnostic

; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Organon Teknika Corporation

; ADDRESSEE: Biotechnology Research Institute

; STREET: 1350-A Piccard Drive

; CITY: Rockville

; STATE: Maryland

; COUNTRY: U.S.A.

; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk


```

;
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/059,853
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/797,554
; FILING DATE: 22-NOV-1991
; APPLICATION NUMBER: US 07/494,991
; FILING DATE: 16-MAR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: William M. Blackstone
; REGISTRATION NUMBER: 29,772
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 258-5200
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3898 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-059-853-2

```

```

Query Match      31.0%; Score 44; DB 2; Length 3898;
Best Local Similarity 70.0%; Pred. No. 2.8e+02;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 6 LVCLIRILRG 15
   |||::|||
Db 663 LVCLIRKVRG 672

```

```

RESULT 5
US-08-976-255-14
; Sequence 14, Application US/08976255
; Patent No. 6136581
; GENERAL INFORMATION:
; APPLICANT: Jono, Keith E.
; TITLE OF INVENTION: KINASE GENES AND USES
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FASTSEQ for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/976,255
; FILING DATE: No. 6136581ember 21, 1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/031,675
; FILING DATE: No. 6136581ember 22, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 229/182
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440

```

```

;
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1503 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; US-08-976-255-14

```

```

Query Match      29.6%; Score 42; DB 4; Length 1503;
Best Local Similarity 35.0%; Pred. No. 2e+02;
Matches 7; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 4 FLVCLIRILRGAMSAKOHV 23
   || : ||::|||:
Db 723 FLVCLIRILRGAMSAKOHV 742

```

```

RESULT 6
5206163-1
; Patent No. 5206163
; APPLICANT: RENARD, ANDRE;DINA, DINO;MARTIAL, JOSEPH
; TITLE OF INVENTION: DNA ENCODING BOVINE DIARRHEA
; VIRUS PROTEIN
; NUMBER OF SEQUENCES: 3
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/550,816
; FILING DATE: 06-JUL-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 331,037
; FILING DATE: 29-MAR-1989
; APPLICATION NUMBER: 752,981
; FILING DATE: 08-JUL-1985
; SEQ ID NO: 1:
; LENGTH: 1286
; 5206163-1

```

```

Query Match      28.9%; Score 41; DB 6; Length 1286;
Best Local Similarity 60.0%; Pred. No. 2.5e+02;
Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 6 LVCLIRILRG 15
   |||::|||
Db 666 LVCLIRKVRG 675

```

```

RESULT 7
US-08-786-748A-38
; Sequence 38, Application US/08786748A
; Patent No. 5714577
; GENERAL INFORMATION:
; APPLICANT: Ronald, Montelaro C.
; APPLICANT: Tencza, Sarah B.
; APPLICANT: Mletzner, Timothy A.
; TITLE OF INVENTION: NOVEL ANTIMICROBIAL PEPTIDES
; NUMBER OF SEQUENCES: 169
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112-0228
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: IBM Compatible
; SOFTWARE: FASTSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/786,748A

```

TOPOLOGY: linear

GENERAL INFORMATION:

```

; APPLICANT: Ronald, Montelaro C.
; APPLICANT: Tencza, Sarah B.
; APPLICANT: Mietzner, Timothy A.
; TITLE OF INVENTION: NOVEL ANTIMICROBIAL PEPTIDES
; NUMBER OF SEQUENCES: 169
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BAKER & BOTTS, L.L.P.
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112-0228
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: FASTSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/932,682
; FILING DATE: 18-SEP-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/786,748
; FILING DATE: 24-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Rochelle K. Seide
; REGISTRATION NUMBER: 32,300
; REFERENCE/DOCKET NUMBER: AP30421-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-705-5000
; TELEFAX: 212-765-2519
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 5945507e
; US-08-932-682-39

Query Match      28.2%; Score 40; DB 2; Length 21;
Best Local Similarity 46.7%; Pred. No. 3.3;
Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 9 LIRIRGAMSAKQHV 23
   :|::|::| 1 :|:
Db 2 VIRVVGACRAIRHI 16

RESULT 11
US-08-452-531-3
; Sequence 3, Application US/08452531
; Patent No. 5733778
; GENERAL INFORMATION:
; APPLICANT: MATLASHESKI, Gregory
; APPLICANT: CHAREST, Hugues
; TITLE OF INVENTION: GENES OF LEISHMANIA WHICH ARE DIFFERENTIALLY-
; TITLE OF INVENTION: EXPRESSED IN AMASTIGOTE FORM
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Slim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/08/452,531
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/115,987
; FILING DATE: 03-SEP-1993
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-448 MIS:dh
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 236 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-452-531-3

Query Match      28.2%; Score 40; DB 1; Length 236;
Best Local Similarity 40.0%; Pred. No. 52;
Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 3 ILLVCLIRIRGAMSAKH 22
   :|::|::| 1 :|:
Db 10 VLLVLCVAVALASAEPH 29

RESULT 12
US-08-460-746A-3
; Sequence 3, Application US/08460746A
; Patent No. 5780591
; GENERAL INFORMATION:
; APPLICANT: MATLASHESKI, Gregory
; APPLICANT: CHAREST, Hugues
; TITLE OF INVENTION: DIFFERENTIALLY EXPRESSED LEISHMANIA
; TITLE OF INVENTION: GENES AND PROTEINS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Slim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,746A
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/302,463
; FILING DATE: 12-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/115,987
; FILING DATE: 03-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-483 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:

```

LENGTH: 236 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-460-746A-3

Query Match 28.2%; Score 40; DB 1; Length 236;
Best Local Similarity 40.0%; Pred. No. 52;
Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 3 IFLLVCLIRLRGMSAKOH 22
:||||: :| :||:|
Db 10 VLLVCVAVALSASAEPH 29

RESULT 13
US-08-460-555-3
Sequence 3, Application US/08460555
Patent No. 5827671
Patent No. 5827671 5827671
GENERAL INFORMATION:
APPLICANT: MATLASHWSKI, Gregory
TITLE OF INVENTION: DIFFERENTIALLY EXPRESSED LEISHMANIA
NUMBER OF SEQUENCES: 4
CURRENT APPLICATION DATA:
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
APPLICATION NUMBER: US/08/460,555
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/302,463
FILING DATE: 12-SEP-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/115,987
FILING DATE: 03-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, Michael I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-485 MIS:V9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 236 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-460-555-3

Query Match 28.2%; Score 40; DB 2; Length 236;
Best Local Similarity 40.0%; Pred. No. 52;
Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;
QY 3 IFLLVCLIRLRGMSAKOH 22
:||||: :| :||:|

Db 10 VLLVCVAVALSASAEPH 29

RESULT 14
US-08-460-066-3
Sequence 3, Application US/08460066
Patent No. 6133017
GENERAL INFORMATION:
APPLICANT: MATLASHWSKI, Gregory
TITLE OF INVENTION: DIFFERENTIALLY EXPRESSED LEISHMANIA
NUMBER OF SEQUENCES: 4
CURRENT APPLICATION DATA:
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
APPLICATION NUMBER: US/08/460,066
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/302,463
FILING DATE: 12-SEP-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/115,987
FILING DATE: 03-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, Michael I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-484 MIS:V9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 236 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-460-066-3

Query Match 28.2%; Score 40; DB 4; Length 236;
Best Local Similarity 40.0%; Pred. No. 52;
Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 3 IFLLVCLIRLRGMSAKOH 22
:||||: :| :||:|
Db 10 VLLVCVAVALSASAEPH 29

RESULT 15
US-08-035-928-2
Sequence 2, Application US/08035928
Patent No. 5538844
GENERAL INFORMATION:
APPLICANT: Duyao, Mabel P.
APPLICANT: Macdonald, Marcy E.
APPLICANT: Gusella, James F.
TITLE OF INVENTION: A No. 5538844el Transport Protein Gene from
the Huntington's Disease Region
NUMBER OF SEQUENCES: 21

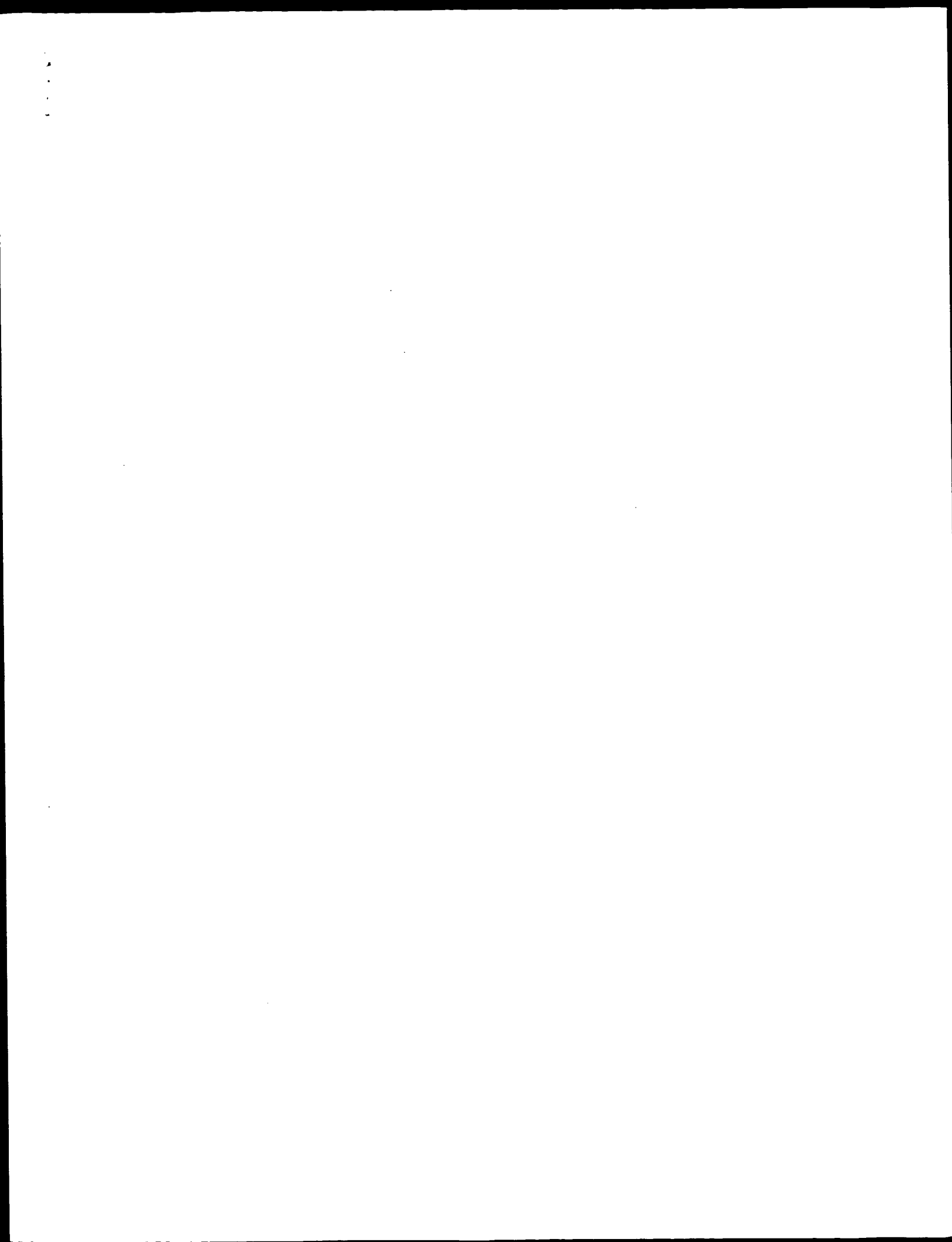
* Thu Jan 10 08:15:58 2002

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Sterne, Kessler, Goldstein & Fox
;; STREET: 1225 Connecticut Avenue N.W.
;; CITY: Washington
;; STATE: D.C.
;; COUNTRY: U.S.A.
;; ZIP: 20036
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/035,928
;; FILING DATE: 19930323
;; CLASSIFICATION: 435
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202) 466-0800
;; TELEFAX: (202) 833-8716
;; TELEX:
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 455 amino acids
;; TYPE: AMINO ACID
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-035-928-2

Query Match 28.2%; Score 40; DB 1; Length 455;
Best Local Similarity 34.6%; Pred. No. 1.1e+02;
Matches 9; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

Qy 4 FLVCLIRIRGMSAKOHVGFEMAA 29
|| : : || : : || : : || : : ||
Db 313 FLIGLTMATIGCAVARIRHPGGEVAA 338

Search completed: January 10, 2002, 02:01:57
Job time: 279 sec



```

2 Ser1epheleuValCysleulleaArgileuAtrgylalameTs
||| ||||| ::|||::||| ||| ::::
171 TCACACTTCTCATCTCTGCTTCAATCCGCCACACTAAATTTCACTTAA
18 rAlaYSgInHsValGlyPheGluMeAlaA 29
::::|||::||| ||||| |||||
121 ATCCAAACATCACTTGGCTTGGAAGCGCGGCC 88

seq_name: /cgn2.6/ptodata/2/ina/6E_COMB.seq:US-09-030-607-E
seq_documentation_block:
; Sequence 6, Application US/09030607
; Patent No. 6262245
GENERAL INFORMATION:

```

APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, David C.
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS FO
NUMBER OF SEQUENCES: 224
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/030,607
FILING DATE: 25-FEB-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: MAKI, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.427C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 682-6031
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 818 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-030-607-6

alignment_scores:
Quality: 64.00 Length: 28
Ratio: 3.200 Gaps: 0
Percent Similarity: 71.429 Percent Identity: 50.000

alignment_block:

US-09-712-768-8 x US-09-030-607-6/rev ..

Align seg 1/1 to reverse of: US-09-030-607-6 from: 1 to: 818

```
2 SerIlePheLeuLeuValCysLeuIleArgIleLeuArgGlyAlaMetse 18
||| ||||| ::|||::||| ||| ::|||
171 TCAACTTTCCTCAGTATGCTTCATCCGCCAACAATAATATTCACCTTAC 122
18 rAlaLysGlnHisValGlyPheGluMetAlaIa 29
:::|||::||| ||||| |||||
121 ATCCAAACATCATCTTGGCTTGAAGCCGCCGCC 88
```

seq_name: /cgn2_6/ptodata/2/1na/5A_COMB.seq:US-08-219-842-3

seq_documentation_block:

Sequence 3, Application US/08219842
Patent No. 5565323
GENERAL INFORMATION:
APPLICANT: Parker, W. D.
APPLICANT: Herrnstadt, Corinna
TITLE OF INVENTION: Diagnostic and Therapeutic Compositions
TITLE OF INVENTION: for Alzheimer's Disease
NUMBER OF SEQUENCES: 95
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/219,842
FILING DATE: 30-MAR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-AG 9504
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 856 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-219-842-3

alignment_scores:
Quality: 64.00 Length: 28
Ratio: 3.200 Gaps: 0
Percent Similarity: 71.429 Percent Identity: 50.000

alignment_block:

US-09-712-768-8 x US-08-219-842-3 ..

Align seg 1/1 to: US-08-219-842-3 from: 1 to: 856

```
2 SerIlePheLeuLeuValCysLeuIleArgIleLeuArgGlyAlaMetse 18
||| ||||| ::|||::||| ||| ::|||
704 TCAACTTTCCTCAGTATGCTTCATCCGCCAACAATAATATTCACCTTAC 753
18 rAlaLysGlnHisValGlyPheGluMetAlaIa 29
:::|||::||| ||||| |||||
754 ATCCAAACATCATCTTGGCTTGAAGCCGCCGCC 787
```

seq_name: /cgn2_6/ptodata/2/1na/5A_COMB.seq:US-08-451-096-3

seq_documentation_block:

Sequence 3, Application US/08451096
Patent No. 5760205
GENERAL INFORMATION:
APPLICANT: Parker, W. D.
APPLICANT: Herrnstadt, Corinna
TITLE OF INVENTION: Diagnostic and Therapeutic Compositions
TITLE OF INVENTION: for Alzheimer's Disease
NUMBER OF SEQUENCES: 95
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/451,096
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/219,842
FILING DATE: 30-MAR-1994


```

; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-AG 9504
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 856 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-451-096-3

alignment_scores:
Quality: 64.00 Length: 28
Ratio: 3.200 Gaps: 0
Percent Similarity: 71.429 Percent Identity: 50.000

alignment_block:
US-09-712-768-8 x US-08-451-096-3 ..

Align seq 1/1 to: US-08-451-096-3 from: 1 to: 856

2 SerTlePheLeuLeuValCysLeuIleArgIleLeuArgGlyAlaMetse 18
||| ||||| ::|||::||| ||| ::|||
704 TCACACTTCTCCTACTATCTGCTTCATCCGCCACATAATATTTCACCTTAC 753
18 rAlaLysGlnHisValGlyPheGluMetAla 29
::|||::||| ||||| |||||
754 ATCCAACATCATCTTGGCTTGCAGAGCCGCCGCC 787

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-810-599-82

seq_documentation_block:
; Sequence 82, Application US/08810599
; Patent No. 5976798
; GENERAL INFORMATION:
; APPLICANT: PARKER, W. Davis
; APPLICANT: HERRNSTADT, Corinna
; APPLICANT: GHOSH, Soumitra S.
; APPLICANT: FAHY, Eoin
; TITLE OF INVENTION: Methods for Detecting Mitochondrial Mutations
; TITLE OF INVENTION: Diagnostic for Alzheimer's Disease and Methods for Determining
; TITLE OF INVENTION: of Mitochondrial Nucleic Acid
; NUMBER OF SEQUENCES: 82
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: 1025 Connecticut Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: US
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.25" floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Wordperfect 6.1 for Windows
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/810,599
; FILING DATE: Concurrent Herewith
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/757,438
; FILING DATE: 27 Nov. 5976798 1996
; APPLICATION NUMBER: US 08/614,072
; FILING DATE: 12 Mar 1996
; APPLICATION NUMBER: US 08/536,036
; FILING DATE: 29 Sep 1995
; APPLICATION NUMBER: US 08/414,969
; FILING DATE: 31 Mar 1995

```

```

; APPLICATION NUMBER: US 08/413,740
; FILING DATE: 30 Mar 1995
; APPLICATION NUMBER: US 08/410,658
; FILING DATE: 24 MARCH 1995
; APPLICATION NUMBER: US 08/397,808
; FILING DATE: 3 Mar 1995
; APPLICATION NUMBER: US 08/219,842
; FILING DATE: 30 MARCH 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Tofteneftl, Judith L.
; REGISTRATION NUMBER: 39,048
; REFERENCE/DOCKET NUMBER: 2105/17
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-429-1776
; TELEFAX: 202-429-0796
; INFORMATION FOR SEQ ID NO: 82:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 926 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; HYPOTHETICAL: No
; ANTI-SENSE: No
US-08-810-599-82

alignment_scores:
Quality: 64.00 Length: 28
Ratio: 3.200 Gaps: 0
Percent Similarity: 71.429 Percent Identity: 50.000

alignment_block:
US-09-712-768-8 x US-08-810-599-82 ..

Align seq 1/1 to: US-08-810-599-82 from: 1 to: 926

2 SerTlePheLeuLeuValCysLeuIleArgIleLeuArgGlyAlaMetse 18
||| ||||| ::|||::||| ||| ::|||
704 TCACACTTCTCCTACTATCTGCTTCATCCGCCACATAATATTTCACCTTAC 753
18 rAlaLysGlnHisValGlyPheGluMetAla 29
::|||::||| ||||| |||||
754 ATCCAACATCATCTTGGCTTGCAGAGCCGCCGCC 787

seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-413-740A-3

seq_documentation_block:
; Sequence 3, Application US/08413740A
; Patent No. 6171859
; GENERAL INFORMATION:
; APPLICANT: HERRNSTADT, CORINNA
; APPLICANT: PARKER, WILLIAM D.
; APPLICANT: DAVIS, ROBERT
; APPLICANT: MILLER, SCOTT W.
; TITLE OF INVENTION: Diagnosis, Therapy and Cellular and
; TITLE OF INVENTION: Animal Models for Diseases Associated with Mitochondrial
; NUMBER OF SEQUENCES: 206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kenyon & Kenyon
; STREET: 1025 Connecticut Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20036-5405
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/413,740A

```

```

:
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/US95/04063
: FILING DATE: 30-MAR-1995
: APPLICATION NUMBER: 08/413,740
: FILING DATE: 30-MAR-1995
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Bonham, David B.
: REGISTRATION NUMBER: 34297
: REFERENCE/DOCKET NUMBER: 2105/7
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202) 429-0796
: TELEFAX: (202) 429-0796
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 954 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: other nucleic acid
: HYPOTHEICAL: NO
: ANTI-SENSE: NO
: US-08-413-740A-3

alignment_scores:
      Quality: 64.00      Length: 28
      Ratio: 3.200      Gaps: 0
      Percent Similarity: 71.429      Percent Identity: 50.000

alignment_block:
US-09-712-768-8 x US-08-413-740A-3  ..

Align seg 1/1 to: US-08-413-740A-3 from: 1 to: 954

2 SerIlePheLeuValCysLeuIleArgIleuArgGlyAlaMetse 18
||| ||||| ::|||::||| ||| ::|||
703 TCACTTTCCTCAGTATCTGCTTCATCCGCCAATATATTCACCTTAC 752
18 rAlaIysGlnHisValGlyPheGluMetAlaIa 29
:::|||||::||| ||||| |||||
753 ATCCAAACATCATCTTGCTTGCATGGAAGCGCGCC 786

seq_name: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US95-04063-3

seq_documentation_block:
: Sequence 3, Application PC/TUS9504063
: GENERAL INFORMATION:
: APPLICANT: HERRNSTADT, CORRINA
: APPLICANT: PARKER, WILLIAM D.
: APPLICANT: DAVIS, ROBERT W.
: APPLICANT: MILLER, SCOTT W.
: TITLE OF INVENTION: Diagnosis, Therapy and Cellular and
: TITLE OF INVENTION: Animal Models for Diseases Associated with Mitochondrial
: TITLE OF INVENTION: Defects
: NUMBER OF SEQUENCES: 206
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Kenyon & Kenyon
: STREET: 1025 Connecticut Avenue, N.W.
: CITY: Washington
: STATE: DC
: COUNTRY: USA
: ZIP: 20036-5405
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US95/04063
: FILING DATE: 30-MAR-1995
```

```

:
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Bonham, David B.
: REGISTRATION NUMBER: 34297
: REFERENCE/DOCKET NUMBER: 2105/7
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202) 429-0796
: TELEFAX: (202) 429-0796
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 954 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: other nucleic acid
: HYPOTHEICAL: NO
: ANTI-SENSE: NO
: PCT-US95-04063-3

alignment_scores:
      Quality: 64.00      Length: 28
      Ratio: 3.200      Gaps: 0
      Percent Similarity: 71.429      Percent Identity: 50.000

alignment_block:
US-09-712-768-8 x PCT-US95-04063-3  ..

Align seg 1/1 to: PCT-US95-04063-3 from: 1 to: 954

2 SerIlePheLeuValCysLeuIleArgIleuArgGlyAlaMetse 18
||| ||||| ::|||::||| ||| ::|||
703 TCACTTTCCTCAGTATCTGCTTCATCCGCCAATATATTCACCTTAC 752
18 rAlaIysGlnHisValGlyPheGluMetAlaIa 29
:::|||||::||| ||||| |||||
753 ATCCAAACATCATCTTGCTTGCATGGAAGCGCGCC 786

seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:US-09-097-889-2

seq_documentation_block:
: Sequence 2, Application US/09097889
: Patent No. 6218117
: GENERAL INFORMATION:
: APPLICANT: HERRNSTADT, CORRINA
: APPLICANT: Ghosh, Soumitra S.
: APPLICANT: Davis, Robert E.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING
: TITLE OF INVENTION: AGENTS THAT QUANTITATIVELY ALTER DETECTABLE
: NUMBER OF INVENTION: EXTRAMITOCHONDRIAL DNA: MITOCHONDRIAL DNA RATIOS
: NUMBER OF SEQUENCES: 26
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SEED and BERRY LLP
: STREET: 6300 Columbia Center, 701 Fifth Avenue
: CITY: Seattle
: STATE: Washington
: COUNTRY: USA
: ZIP: 98104
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/097,889
: FILING DATE: 15-JUN-1998
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Roseman Ph.D., Stephen J.
: REGISTRATION NUMBER: 43,058
: REFERENCE/DOCKET NUMBER: 660088.417
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (206) 622-4900
```

TELEFAX: (206) 682-6031
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 16569 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-09-097-889-2

alignment_scores:
 Quality: 64.00 Length: 28
 Ratio: 3.200 Gaps: 0
 Percent Similarity: 71.429 Percent Identity: 50.000

alignment_block:
 US-09-712-768-8 x US-09-097-889-2 ..

Align seg 1/1 to: US-09-097-889-2 from: 1 to: 16569

2 SerIephLeuValCysLeuIleArgIleLeuArgIlyAlaMetse 18
 ||| ||||| ::|||::||| |||
 9840 TCACACTTCTCTCACTATCTCTTCATCCGCCAACATATTTTCACTTTAC 9889
 18 rAlaIysGlnHisValGlyPheGluMetAlaAla 29
 ::|||::||| ||||| |||||
 9890 ATCCAAACATCACTTGTGCTTCAGAGCCGCCGCC 9923

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-920-812-8

seq_documentation_block:
 ; Sequence 8, Application US/08920812
 ; Patent No. 5763188
 ; GENERAL INFORMATION:
 ; APPLICANT: Ohno, Tsuneya
 ; APPLICANT: Matsuhisa, Akio
 ; APPLICANT: Uehara, Hirotsugu
 ; APPLICANT: Eda, Soji
 ; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
 ; NUMBER OF SEQUENCES: 25
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 ; STREET: 6300 Sears Tower, 233 South Wacker Drive
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: United States of America
 ; ZIP: 60606-6402
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/920,812
 ; FILING DATE: 29-AUG-1997
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/362,577
 ; FILING DATE: 27-MAR-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Rin-Laures, Li-Hsien
 ; REGISTRATION NUMBER: 33,547
 ; REFERENCE/DOCKET NUMBER: 19036/32420
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312/474-6300
 ; TELEFAX: 312/474-0448
 ; TELEX: 25-3856
 ; INFORMATION FOR SEQ ID NO: 8:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3287 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear

MOLECULE TYPE: Genomic DNA
 ORIGINAL SOURCE:
 ORGANISM: Staphylococcus epidermidis
 STRAIN: Clinical Isolate SE-37
 US-08-920-812-8

alignment_scores:
 Quality: 58.00 Length: 25
 Ratio: 2.900 Gaps: 0
 Percent Similarity: 80.000 Percent Identity: 40.000

alignment_block:
 US-09-712-768-8 x US-08-920-812-8 ..

Align seg 1/1 to: US-08-920-812-8 from: 1 to: 3287

5 LeuLeuValCysLeuIleArgIleLeuArgIlyAlaMetSerAlaIysG1 21
 ||| ||||| ||||| ::|||::||| |||
 2607 CTACTGTTGTCTAGTCAATTAATTAAGTCCCACTACTGCTGATAG 2656
 21 nHisValGlyPheGluMetAlaAla 29
 : |||::|||:: :
 2657 AGTTGTATCATTTGATGCTCGAGT 2681

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-920-827-8

seq_documentation_block:
 ; Sequence 8, Application US/08920827
 ; Patent No. 5770375
 ; GENERAL INFORMATION:
 ; APPLICANT: Ohno, Tsuneya
 ; APPLICANT: Matsuhisa, Akio
 ; APPLICANT: Uehara, Hirotsugu
 ; APPLICANT: Eda, Soji
 ; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
 ; NUMBER OF SEQUENCES: 25
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 ; STREET: 6300 Sears Tower, 233 South Wacker Drive
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: United States of America
 ; ZIP: 60606-6402
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/920,827
 ; FILING DATE: 29-AUG-1997
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/362,577
 ; FILING DATE: 27-MAR-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Rin-Laures, Li-Hsien
 ; REGISTRATION NUMBER: 33,547
 ; REFERENCE/DOCKET NUMBER: 19036/32420
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312/474-6300
 ; TELEFAX: 312/474-0448
 ; TELEX: 25-3856
 ; INFORMATION FOR SEQ ID NO: 8:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3287 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: Genomic DNA
 ; ORIGINAL SOURCE:
 ; ORGANISM: Staphylococcus epidermidis

STRAIN: Clinical Isolate SE-37
US-08-920-827-8

alignment_scores:
Quality: 58.00 Length: 25
Ratio: 2.900 Gaps: 0
Percent Similarity: 80.000 Percent Identity: 40.000

alignment_block:
US-09-712-768-8 x US-08-920-827-8 ..

Align seg 1/1 to: US-08-920-827-8 from: 1 to: 3287

5 LeuLeuValCysLeuIleArgIleLeuArgGlyAlaMetSerAlaLysG1 21
|||||
2607 CTACTGTTGTGTCTAGTCAGATTAAAGGTCACACTGCTGATAG 2656
21 nh1sValGlyPheGluMetAlaIa1a 29
: |||
2657 AGTTGTATCATTTGATGCTCGAGT 2681

seq_name: /cgn2_6/prodata/2/lna/5A_COMB.seq:US-08-921-177-8

seq_documentation_block:
Sequence 8, Application US/08921177
Patent No. 5798211
GENERAL INFORMATION:
APPLICANT: Ohno, Tsuneya
APPLICANT: Matsuhisa, Akio
APPLICANT: Uehara, Hirotsugu
APPLICANT: Eda, Soji
TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/921.177
FILING DATE: 29-AUG-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,577
FILING DATE: 27-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 19036/32420
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 3287 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
ORIGINAL SOURCE:
ORGANISM: Staphylococcus epidermidis
STRAIN: Clinical Isolate SE-37
US-08-921-177-8

alignment_scores:
Quality: 58.00 Length: 25
Ratio: 2.900 Gaps: 0
Percent Similarity: 80.000 Percent Identity: 40.000

alignment_block:
US-09-712-768-8 x US-08-921-177-8 ..

Align seg 1/1 to: US-08-921-177-8 from: 1 to: 3287

5 LeuLeuValCysLeuIleArgIleLeuArgGlyAlaMetSerAlaLysG1 21
|||||
2607 CTACTGTTGTGTCTAGTCAGATTAAAGGTCACACTGCTGATAG 2656
21 nh1sValGlyPheGluMetAlaIa1a 29
: |||
2657 AGTTGTATCATTTGATGCTCGAGT 2681

seq_name: /cgn2_6/prodata/2/lna/5A_COMB.seq:US-08-362-577C-8

seq_documentation_block:
Sequence 8, Application US/08362577C
Patent No. 5807673
GENERAL INFORMATION:
APPLICANT: Ohno, Tsuneya
APPLICANT: Matsuhisa, Akio
APPLICANT: Uehara, Hirotsugu
APPLICANT: Eda, Soji
TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/362,577C
FILING DATE: 27-MAR-1995
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 19036/32420
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 3287 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
ORIGINAL SOURCE:
ORGANISM: Staphylococcus epidermidis
STRAIN: Clinical Isolate SE-37
US-08-362-577C-8

alignment_scores:
Quality: 58.00 Length: 25
Ratio: 2.900 Gaps: 0
Percent Similarity: 80.000 Percent Identity: 40.000

alignment_block:
US-09-712-768-8 x US-08-362-577C-8 ..

Align seg 1/1 to: US-08-362-577C-8 from: 1 to: 3287

5 LeuLeuValCysLeuIleAArgIleLeuAArgGlyAlaMetSerAlaIysG1 21
|||||
2607 CTAAGTGTGTTGCTAGTACGATTAATTAAGGTCACACTACTGCTGATAG 2656
21 nhlsva1g1yphg1umeta1a1a 29
: |||::|||::: :::::
2657 AGTTGATCATTTGATGCTCGAGT 2681

seq_name: /cgn2_6/ptodata/2/1na/5B_COMB.seq:US-08-920-828-8

seq_documentation_block:

Sequence 8, Application US/08920828
Patent No. 5853998
GENERAL INFORMATION:
APPLICANT: Ohno, Tsuneya
APPLICANT: Matsuhisa, Akio
APPLICANT: Uehara, Hirotsugu
APPLICANT: Ega, Soji
TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/920.828
FILING DATE: 29-AUG-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,577
FILING DATE: 27-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 19036/32420
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 3287 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
ORIGINAL SOURCE:
ORGANISM: staphylococcus epidermidis
STRAIN: clinical isolate SE-37
US-08-920-828-8

alignment_scores:

Quality: 58.00 Length: 25
Ratio: 2.900 Gaps: 0
Percent Similarity: 80.000 Percent Identity: 40.000

alignment_block:

US-09-712-768-8 x US-08-920-828-8 ..

Align seg 1/1 to: US-08-920-828-8 from: 1 to: 3287

5 LeuLeuValCysLeuIleAArgIleLeuAArgGlyAlaMetSerAlaIysG1 21
|||||
2607 CTAAGTGTGTTGCTAGTACGATTAATTAAGGTCACACTACTGCTGATAG 2656
21 nhlsva1g1yphg1umeta1a1a 29
: |||::|||::: :::::
2657 AGTTGATCATTTGATGCTCGAGT 2681

seq_name: /cgn2_6/ptodata/2/1na/5A_COMB.seq:US-08-484-105-7

seq_documentation_block:

Sequence 7, Application US/08484105.
Patent No. 585341
GENERAL INFORMATION:
APPLICANT: STILLMAN, Bruce
APPLICANT: BELL, Stephen P
APPLICANT: KOBAYASHI, Ryuji
APPLICANT: RINE, Jasper
APPLICANT: FOSS, Margit
APPLICANT: McNALLY, Francis J
APPLICANT: LAURENSEN, Patricia
APPLICANT: HERSKOWITZ, Ira
APPLICANT: LI, Joachim J
TITLE OF INVENTION: Kimberly
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,105
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Osman Ph.D., Richard Aron
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-59032/DJB/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 494-8700
TELEFAX: (415) 494-8771
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2404 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-484-105-7

alignment_scores:

Quality: 49.00 Length: 24
Ratio: 2.450 Gaps: 0
Percent Similarity: 83.333 Percent Identity: 41.667

alignment_block:

US-09-712-768-8 x US-08-484-105-7/rev ..

Align seg 1/1 to reverse of: US-08-484-105-7 from: 1 to: 2404

2 SerIlePheLeuLeuValCysLeuIleAArgIleLeuAArgGlyAlaMetSer 18

```
||| |||:|||||:|||||:|||||: |||: |||
1741 TCGAATTCATCTAATATGCTCTATTCGAATTCATCGAGGTCGATAG 1692
18 rAlaLysGlnHisValGlyPhe 25
|::: |||:|||||
1691 TTCTTTTCCACGCTTCATTC 1670
```

seq_name: /cgn2_6/prodata/2/lna/5A_COMB.seq:US-08-484-106-7

```
seq_documentation_block:
; Sequence 7, Application US/08484106
; Patent No. 5614618
; GENERAL INFORMATION:
; APPLICANT: STILLMAN, Bruce
; APPLICANT: BELL, Stephen P
; APPLICANT: KOBAYASHI, Ryuji
; APPLICANT: RINE, Jasper
; APPLICANT: FOSS, Margit
; APPLICANT: MCNALLY, Francis J
; APPLICANT: LAURENSEN, Patricia
; APPLICANT: HERSKOWITZ, Ira
; APPLICANT: LI, Joachim J
; APPLICANT: GAVIN, Kimberly
; TITLE OF INVENTION: ORIGIN OF REPLICATION COMPLEX GENES
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,106
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman Ph.D., Richard Aron
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: A-59032/DJB/RAO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 494-8700
; TELEFAX: (415) 494-8771
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2404 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-484-106-7
```

```
alignment_scores:
Quality: 49.00 Length: 24
Ratio: 2.450 Gaps: 0
Percent Similarity: 83.333 Percent Identity: 41.667
```

alignment_block:

US-09-712-768-8 x US-08-484-106-7/rev ..

Align seg 1/1 to reverse of: US-08-484-106-7 from: 1 to: 2404

```
2 SerIlePheLeuLeuValCysLeuIleArgIleLeuArgGlyAlaMetCse 18
||| |||:|||||:|||||:|||||: |||: |||
1741 TCGAATTCATCTAATATGCTCTATTCGAATTCATCGAGGTCGATAG 1692
```

```
18 rAlaLysGlnHisValGlyPhe 25
|::: |||:|||||
1691 TTCTTTTCCACGCTTCATTC 1670
```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2002, 02:00:24 ; Search time 18.96 seconds

(without alignments)
116.512 Million cell updates/sec

Title: US-09-712-768-8

Perfect score: 142

Sequence: 1 DSIFLVCLIRILRGAMSAKQHGFEEMA 29

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	105	73.9	274	2	S03807
2	89	62.7	266	2	C45164
3	79	55.6	281	2	T11235
4	78	54.9	261	2	T11330
5	78	54.9	265	1	OTWT3M
6	77	54.2	272	2	F58930
7	75	52.8	261	1	OTMS3
8	75	52.8	261	2	S55010
9	75	52.8	261	2	S04753
10	74	52.1	265	1	OTR23M
11	74	52.1	265	1	S20801
12	73	51.4	260	2	T11798
13	73	51.4	261	2	T11199
14	73	51.4	265	2	A48304
15	71	50.0	262	1	OTFF3
16	71	50.0	262	1	OTFF3Y
17	70	49.3	261	2	T11186
18	70	49.3	261	2	T11146
19	70	49.3	261	2	G59153
20	69	48.6	261	2	B40076
21	69	48.6	261	2	T10992
22	69	48.6	261	2	T11826
23	69	48.6	261	2	T11839
24	69	48.6	261	2	T11028
25	69	48.6	261	2	T11469
26	69	48.6	265	1	OTSY3M
27	69	48.6	272	2	S43951
28	68	47.9	260	2	H34284
29	68	47.9	261	2	T11525

ALIGNMENTS

RESULT 1

S03807

Cytochrome-c oxidase (EC 1.9.3.1) chain III - Paracoccus denitrificans

C:Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 11-Jun-1999

C:Accession: S03807; S02524; A24371

R:Raitio, M.; Jalli, T.; Saraste, M.

EMBO J. 6, 2825-2833, 1987

A:Title: Isolation and analysis of the genes for cytochrome c oxidase in Paracoccus d

A:Reference number: S03803

A:Accession: S03807

A:Molecule type: DNA

A:Residues: 1-274 <RA1>

A:Cross-references: EMBL:X05828; NID:945468; PIDN:CAA29272.1; PID:945473

R:Raitio, M.; Jalli, T.; Puustinen, A.; Finel, M.

Eur. J. Biochem. 172, 543-546, 1988

A:Title: The Paracoccus denitrificans cytochrome aa(3) has a third subunit.

A:Reference number: S02524; MUID:8616732

A:Accession: S02524

A:Molecule type: protein

A:Residues: 2-16 <HAL>

R:Saraste, M.; Raitio, M.; Peramaa, A.

FEBS Lett. 206, 154-156, 1986

A:Title: A gene in Paracoccus for subunit III of cytochrome oxidase.

A:Reference number: A24371; MUID:87005242

A:Accession: A24371

A:Molecule type: DNA

A:Residues: 1-144 <SAR>

A:Cross-references: GB:X04406; NID:945466; PIDN:CAA27995.1; PID:945467

C:Genetics:

A:Gene: COIII

C:Superfamily: cytochrome-c oxidase chain III

C:Keywords: electron transfer; membrane-associated complex; oxidative phosphorylation

F:2-274/Product: cytochrome-c oxidase chain III #status experimental <MAN>

Query Match 73.9%; Score 105; DB 2; Length 274;

Best Local Similarity 71.4%; Pred. No. 1.3e-08;

Matches 20; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 DSIFLVCLIRILRGAMSAKQHGFEEMA 29

DB 225 DSIFLVCLIRILRGAMSAKQHGFEEMA 252

RESULT 2

C45164

Cytochrome-c oxidase (EC 1.9.3.1) chain III - Rhodobacter sphaeroides

C:Species: Rhodobacter sphaeroides

C:Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 17-Mar-1999

C:Accession: C45164

R:cao, J.; Hostler, J.; Shapleigh, J.; Revzin, A.; Ferguson-Miller, S.

A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-272 <ARN>
 A:Cross-references: GB:D89861; NID:g4115761; PIDN:BAA34657.1; PID:g3927869
 C:Genetics:
 A:Gene: cox3
 A:Genome: mitochondrion
 C:Superfamily: cytochrome-c oxidase chain III
 C:Keywords: electron transfer; membrane-associated complex; mitochondrial inner membrane

Query Match 54.2%; Score 77; DB 2; Length 272;
 Best Local Similarity 50.0%; Pred. No. 0.00027;
 Matches 14; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 2 SFLLVCLIRLRGAMSAKOHVFEMAA 29
 DB 222 SFLLVCLIRLRGAMSAKOHVFEMAA 249

RESULT 7
 OTMS3
 Cytochrome-c oxidase (EC 1.9.3.1) chain III - mouse mitochondrion
 C:Species: mitochondrion Mus musculus (house mouse)
 C:Date: 02-Apr-1982 #sequence_revision 02-Apr-1982 #text_change 07-Dec-1999
 C:Accession: A00484
 R:Bibb, M.J.; Van Etten, R.A.; Wright, C.T.; Walberg, M.W.; Clayton, D.A.
 Cell 26, 167-180, 1981
 A:Title: Sequence and gene organization of mouse mitochondrial DNA.
 A:Reference number: A00153; MUID:82137051
 A:Accession: A00484
 A:Molecule type: DNA
 A:Residues: 1-261 <BIB>
 A:Cross-references: GB:J01420; NID:g342520; PIDN:AAB48650.1; PID:g896297; EMBL:V00711; N
 C:Genetics:
 A:Gene: coxi1
 A:Genome: mitochondrion
 A:Genetic code: SGC1
 C:Superfamily: cytochrome-c oxidase chain III
 C:Keywords: electron transfer; membrane-associated complex; mitochondrial inner membrane
 F:16-34/Domain: transmembrane helix #status predicted <TR01>
 F:41-66/Domain: transmembrane helix #status predicted <TR02>
 F:73-105/Domain: transmembrane helix #status predicted <TR03>
 F:129-152/Domain: transmembrane helix #status predicted <TR04>
 F:156-183/Domain: transmembrane helix #status predicted <TR05>
 F:191-223/Domain: transmembrane helix #status predicted <TR06>
 F:233-256/Domain: transmembrane helix #status predicted <TR07>

Query Match 52.8%; Score 75; DB 1; Length 261;
 Best Local Similarity 53.6%; Pred. No. 0.00053;
 Matches 15; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 2 SFLLVCLIRLRGAMSAKOHVFEMAA 29
 DB 212 SFLLVCLIRLRGAMSAKOHVFEMAA 239

RESULT 8
 SS5010
 Cytochrome-c oxidase (EC 1.9.3.1) chain III - sea lamprey mitochondrion
 C:Species: mitochondrion Petromyzon marinus (sea lamprey)
 C:Date: 23-Aug-1995 #sequence_revision 03-Nov-1995 #text_change 07-Dec-1999
 C:Accession: SS5010
 R:Lee, W.J.; Kocher, T.D.
 Genetics 139, 873-887, 1995
 A:Title: Complete sequence of a sea lamprey (Petromyzon marinus) mitochondrial genome: e
 A:Reference number: SS4999; MUID:95229067
 A:Accession: SS5010
 A:Molecule type: DNA
 A:Residues: 1-261 <LEE>
 A:Cross-references: EMBL:U11880; NID:g515484; PIDN:AAB08744.1; PID:g515492
 C:Genetics:

A:Gene: coIII
 A:Genome: mitochondrion
 A:Genetic code: SGC1
 C:Superfamily: cytochrome-c oxidase chain III
 C:Keywords: electron transfer; membrane-associated complex; mitochondrial inner membr

Query Match 52.8%; Score 75; DB 2; Length 261;
 Best Local Similarity 53.6%; Pred. No. 0.00053;
 Matches 15; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 2 SFLLVCLIRLRGAMSAKOHVFEMAA 29
 DB 212 SFLLVCLIRLRGAMSAKOHVFEMAA 239

RESULT 9
 S04753
 Cytochrome-c oxidase (EC 1.9.3.1) chain III - rat mitochondrion
 C:Species: mitochondrion Rattus norvegicus (Norway rat)
 C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 07-Dec-1999
 C:Accession: S04753; S65372
 R:Gadaleta, G.; Pepe, G.; De Candia, G.; Quagliariello, C.; Sbisa, E.; Saccone, C.
 J. Mol. Evol. 28, 497-516, 1989
 A:Title: The complete nucleotide sequence of the Rattus norvegicus mitochondrial geno
 A:Reference number: S04747; MUID:89362487
 A:Accession: S04753
 A:Molecule type: DNA
 A:Residues: 1-261 <GAD>
 A:Cross-references: EMBL:X14848; NID:g854269; PIDN:CAA32960.1; PID:g639981
 R:Schaeffer, H.; Noack, H.; Halangk, W.; Brandt, U.; von Jagow, G.
 Eur. J. Biochem. 230, 235-241, 1995
 A:Title: Cytochrome-c oxidase in developing rat heart. Enzymic properties and amino-t
 A:Reference number: S65372; MUID:95324529
 A:Accession: S65372
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 75, 3-11 <SCH>
 C:Genetics:
 A:Genome: mitochondrion
 A:Genetic code: SGC1
 C:Superfamily: cytochrome-c oxidase chain III
 C:Keywords: electron transfer; membrane-associated complex; mitochondrial inner membr

Query Match 52.8%; Score 75; DB 2; Length 261;
 Best Local Similarity 53.6%; Pred. No. 0.00053;
 Matches 15; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 2 SFLLVCLIRLRGAMSAKOHVFEMAA 29
 DB 212 SFLLVCLIRLRGAMSAKOHVFEMAA 239

RESULT 10
 OTR23M
 Cytochrome-c oxidase (EC 1.9.3.1) chain III - rice mitochondrion
 C:Species: mitochondrion Oryza sativa (rice)
 C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 20-Aug-1999
 C:Accession: J00166; S29748; S07875
 R:Kaleikau, E.K.; Andre, C.P.; Walbot, V.
 Nucleic Acids Res. 18, 371, 1990
 A:Title: Sequence of the rice mitochondrial gene for cytochrome oxidase subunit 3.
 A:Reference number: J00166; MUID:90221829
 A:Accession: J00166
 A:Molecule type: DNA
 A:Residues: 1-265 <KAL>
 A:Cross-references: EMBL:X17040; NID:g13216; PIDN:CAA34898.1; PID:g13217
 R:Li, A.W.; Narayanan, K.K.; Andre, C.P.; Kaleikau, E.K.; Walbot, V.
 Curr. Genet. 21, 507-513, 1992
 A:Title: Co-transcription of orf25 and coxIII in rice mitochondria.
 A:Reference number: S26876; MUID:92315351
 A:Accession: S29748

A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-265 <L1U>
 A:Cross-references: EMBL:M74241; NID:g343214; PIDN:AAA66048.1; PID:g804815
 C:Genetics:
 A:Gene: cox3
 A:Genome: mitochondrion
 C:Superfamily: cytochrome-c oxidase chain III
 C:Keywords: electron transfer; membrane-associated complex; mitochondrial inner membrane
 F:1-35/Domain: transmembrane helix #status predicted <TR01>
 F:44-69/Domain: transmembrane helix #status predicted <TR02>
 F:76-108/Domain: transmembrane helix #status predicted <TR03>
 F:132-155/Domain: transmembrane helix #status predicted <TR04>
 F:159-186/Domain: transmembrane helix #status predicted <TR05>
 F:194-226/Domain: transmembrane helix #status predicted <TR06>
 F:236-259/Domain: transmembrane helix #status predicted <TR07>

Query Match 52.1%; Score 74; DB 1; Length 265;
 Best Local Similarity 53.6%; Pred. No. 0.00076;
 Matches 15; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 2 SIFLVCLIRLRGAMSAKOHVFEMAA 29
 ::::||||| |::| ||||| ||
 Db 215 TFLVCGIRGYLGRKHHVGFEMAA 242

RESULT 11
 cytochrome-c oxidase (EC 1.9.3.1) chain III - maize mitochondrion
 C:Species: mitochondrion Zea mays (maize)
 C:Date: 22-Nov-1993 #sequence_revision 26-May-1995 #text_change 20-Aug-1999
 C:Accession: S20801
 R:Kalelkar, E.K.; Malhot, V.
 submitted to the EMBL Data Library, May 1990
 A:Description: Allelic sequence of the maize mitochondrial gene for cytochrome oxidase
 A:Reference number: S20801
 A:Accession: S20801
 A:Molecule type: DNA
 A:Residues: 1-265 <KAL>
 A:Cross-references: EMBL:X53055; NID:g14295; PIDN:CAA37222.1; PID:g14296
 A:Experimental source: strain B37N; seedling
 C:Genetics:
 A:Gene: coxIII
 A:Genome: mitochondrion
 C:Superfamily: cytochrome-c oxidase chain III
 C:Keywords: electron transfer; membrane-associated complex; mitochondrial inner membrane
 F:2-16/Domain: mitochondrial matrix #status predicted <TM01>
 F:17-35/Domain: transmembrane #status predicted <TM02>
 F:36-43/Domain: intracristal #status predicted <ITC1>
 F:44-69/Domain: transmembrane #status predicted <TM02>
 F:70-75/Domain: transmembrane #status predicted <TM02>
 F:76-108/Domain: transmembrane #status predicted <TM03>
 F:109-131/Domain: intracristal #status predicted <ITC2>
 F:132-155/Domain: transmembrane #status predicted <TM04>
 F:156-158/Domain: mitochondrial matrix #status predicted <MM3>
 F:159-186/Domain: transmembrane #status predicted <TM05>
 F:187-193/Domain: intracristal #status predicted <ITC3>
 F:194-226/Domain: transmembrane #status predicted <TM06>
 F:227-235/Domain: mitochondrial matrix #status predicted <MM4>
 F:236-259/Domain: transmembrane #status predicted <TM07>
 F:260-264/Domain: intracristal #status predicted <ITC4>

RESULT 12
 cytochrome-c oxidase (EC 1.9.3.1) chain III - sea urchin (Arbacia lixula) mitochondrion
 C:Species: mitochondrion Arbacia lixula (black urchin)
 C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
 C:Accession: T11798; D39746; PH0855
 R:De Giorgi, C.; Martiradonna, A.; Lanave, C.; Saccocc, C.
 Mol. Phylogenet. Evol. 5, 323-332, 1996
 A:Title: Complete sequence of the mitochondrial DNA in the sea urchin Arbacia lixula:
 A:Reference number: T11798; MUID:96292527
 A:Accession: T11798
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-260 <DEG1>
 A:Cross-references: EMBL:X80396; NID:g1321876; PIDN:CAA56613.1
 A:Experimental source: eggs
 R:De Giorgi, C.; Lanave, C.; Musci, M.D.; Saccocc, C.
 Mol. Biol. Evol. 8, 515-529, 1991
 A:Title: Mitochondrial DNA in the sea urchin Arbacia lixula: evolutionary inferences
 A:Reference number: A39746; MUID:92017217
 A:Accession: D39746
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 152-208, 'NS', 211-260 <DEG2>
 A:Cross-references: EMBL:X53727; NID:g336368; PIDN:AAA98044.1; PID:g1280181
 R:De Giorgi, C.; De Luca, F.; Saccocc, C.
 Gene 103, 249-252, 1991
 A:Title: Mitochondrial DNA in the sea urchin Arbacia lixula: nucleotide sequence diff
 A:Reference number: PH0854; MUID:91365253
 A:Accession: PH0855
 A:Molecule type: DNA
 A:Residues: 157-210, 'H' <DEG3>
 A:Cross-references: GB:M79455; GB:M35442; NID:g336366; PIDN:AAA31638.2; PID:g7304846
 C:Genetics:
 A:Genome: mitochondrion
 A:Genetic code: SGC8
 C:Superfamily: cytochrome-c oxidase chain III
 C:Keywords: electron transfer; membrane-associated complex; mitochondrial inner membr

Query Match 51.4%; Score 73; DB 2; Length 260;
 Best Local Similarity 57.7%; Pred. No. 0.0011;
 Matches 15; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 4 FLLVCLIRLRGAMSAKOHVFEMAA 29
 ||||| |::| || ||||| ||
 Db 213 FLLVCLIRLRGAMSAKOHVFEMAA 238

RESULT 13
 cytochrome-c oxidase (EC 1.9.3.1) chain III - Vidua chalybeata mitochondrion
 C:Species: mitochondrion Vidua chalybeata
 C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 07-Dec-1999
 C:Accession: T11199
 R:Minidell, D.P.; Sorenson, M.D.; Dimcheff, D.E.
 Proc. Natl. Acad. Sci. U.S.A. 95, 10693-10697, 1998
 A:Title: Multiple independent origins of mitochondrial gene order in birds.
 A:Reference number: Z17242
 A:Accession: T11199
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-261 <MIN>
 A:Cross-references: EMBL:AF090341; NID:g4894501; PID:g4894508; PIDN:AA032531.1
 C:Genetics:
 A:Genome: mitochondrion
 A:Genetic code: SGC1
 C:Superfamily: cytochrome-c oxidase chain III
 C:Keywords: electron transfer; membrane-associated complex; mitochondrial inner membr

Query Match 51.4%; Score 73; DB 2; Length 261;
 Best Local Similarity 50.0%; Pred. No. 0.0011;

Matches 14; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 2 SIFLVCLIRIRLGAMSAKOHVGFEMAA 29
 | |||||:|:| | ||| ||
 Db 212 STFLVCLIRLRILKIVHFFPSHHFGFEAAA 239

RESULT 14

A48304
 cytochrome-c oxidase (EC 1.9.3.1) chain III - fava bean mitochondrion
 C:Species: mitochondrion Vicia faba (fava bean)
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 20-Aug-1999
 C:Accession: A48304
 R:Macfarlane, J.L.; Wahlsteiner, J.A.; Wolstenholme, D.R.
 Curr. Genet. 17, 33-40, 1990
 A:Title: A gene for cytochrome c oxidase subunit III (COXIII) in broad bean mitochondria
 A:Reference number: A48304; MUID:90182727
 A:Accession: A48304
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-265 <MAC>
 A:Cross-references: GB:X51690
 C:Genetics:
 A:Genome: mitochondrion
 C:Superfamily: cytochrome-c oxidase chain III
 C:Keywords: electron transfer; membrane-associated complex; mitochondrial inner membrane

Query Match 51.4%; Score 73; DB 2; Length 265;
 Best Local Similarity 50.0%; Pred. No. 0.0011;
 Matches 14; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 2 SIFLVCLIRIRLGAMSAKOHVGFEMAA 29
 :|||:| | | | :|:| ||||| ||
 Db 215 TFLVLCIGIRQYLQGMTEHHVGFEMAA 242

RESULT 15

OTPF3

cytochrome-c oxidase (EC 1.9.3.1) chain III - fruit fly (Drosophila melanogaster) mitoch
 C:Species: mitochondrion Drosophila melanogaster
 C:Date: 18-Apr-1984 #sequence_revision 18-Apr-1984 #text_change 07-Dec-1999
 C:Accession: A00485; A93463; S02250
 R:de Bruijn, M.H.L.
 Nature 304, 234-241, 1983
 A:Title: Drosophila melanogaster mitochondrial DNA, a novel organization and genetic cod
 A:Reference number: A93307; MUID:83245048
 A:Accession: A00485
 A:Molecule type: DNA
 A:Residues: 1-179 <DEB>
 R:Clary, D.O.; Wahlsteiner, J.A.; Wolstenholme, D.R.
 Nucleic Acids Res. 11, 2411-2425, 1983
 A:Title: Transfer RNA genes in Drosophila mitochondrial DNA: related 5' flanking sequenc
 A:Reference number: A93463; MUID:83220794
 A:Accession: A93463
 A:Molecule type: DNA
 A:Residues: 179-262 <CIA>
 R:Garesse, R.
 Genetics 118, 649-663, 1988
 A:Title: Drosophila melanogaster mitochondrial DNA: gene organization and evolutionary c
 A:Reference number: S01185; MUID:88212147
 A:Accession: S02250
 A:Molecule type: DNA
 A:Residues: 179-262 <GAR>
 A:Cross-references: GB:M37275; EMBL:Y00610; NID:q336819; PIDN:AAA69708.1; PID:q336820
 C:Genetics:
 A:Gene: coIII
 A:Cross-references: FLYBase:FBgn0013676
 A:Genome: mitochondrion
 A:Genetic code: SGC4
 C:Superfamily: cytochrome-c oxidase chain III
 C:Keywords: electron transfer; membrane-associated complex; mitochondrial inner membrane
 F:17-35/Domain: transmembrane helix #status predicted <TR01>

F:42-67/Domain: transmembrane helix #status predicted <TR02>
 F:74-106/Domain: transmembrane helix #status predicted <TR03>
 F:130-153/Domain: transmembrane helix #status predicted <TR04>
 F:157-184/Domain: transmembrane helix #status predicted <TR05>
 F:192-224/Domain: transmembrane helix #status predicted <TR06>
 F:234-257/Domain: transmembrane helix #status predicted <TR07>

Query Match 50.0%; Score 71; DB 1; Length 262;
 Best Local Similarity 57.7%; Pred. No. 0.0022;
 Matches 15; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 4 FLVCLIRIRLGAMSAKOHVGFEMAA 29
 |||||:| | | | ||| ||
 Db 215 FLVCLIRLRILNHPKNIHFGFEAAA 240

Search completed: January 10, 2002, 02:00:24
 Job time: 547 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2002, 02:03:12 ; Search time 16.73 Seconds
(without alignments)
63.555 Million cell updates/sec

Title: US-09-712-768-8
142

Perfect score: 1 DSIFLVCLIRILGAMSAKOHVFEMAA 29

Sequence:

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	ID	Description
1	105	73.9	273	1	COX3_PARDE
2	78	54.9	265	1	COX3_AEGCO
3	78	54.9	265	1	COX3_WHEAT
4	77	54.2	147	1	COX3_SPOFR
5	75	52.8	259	1	COX3_LOLBI
6	75	52.8	260	1	COX3_RAT
7	75	52.8	261	1	COX3_MOUSE
8	75	52.8	261	1	COX3_PETMA
9	74	52.1	265	1	COX3_MAIZE
10	74	52.1	265	1	COX3_ORYSA
11	73	51.4	109	1	COX3_ARBLI
12	73	51.4	265	1	COX3_VICFA
13	71	50.0	262	1	COX3_DROME
14	71	50.0	262	1	COX3_DROTA
15	69	48.6	261	1	COX3_GADMO
16	69	48.6	261	1	COX3_HYLLA
17	69	48.6	263	1	COX3_LOCM1
18	69	48.6	265	1	COX3_ARATH
19	69	48.6	265	1	COX3_SOYBN
20	69	48.6	272	1	COX3_CHOCR
21	68	47.9	261	1	COX3_PARLI
22	68	47.9	261	1	COX3_PELSU
23	68	47.9	261	1	COX3_SOUAC
24	68	47.9	261	1	COX3_STRCA
25	68	47.9	265	1	COX3_MARPO
26	68	47.9	265	1	COX3_MYTED
27	68	47.9	272	1	COX3_PYLIL
28	67	47.2	260	1	COX3_XENLA
29	67	47.2	261	1	COX3_CHICK
30	67	47.2	261	1	COX3_PHOVI
31	67	47.2	261	1	COX3_PIG
32	67	47.2	262	1	COX3_PROMI
33	66	46.5	261	1	COX3_CERSI

34	66	46.5	261	1	COX3_EOVAS	P92481 equus asinu
35	66	46.5	261	1	COX3_FELCA	P48892 felis silve
36	66	46.5	261	1	COX3_HORSE	P48661 equus cabal
37	66	46.5	261	1	COX3_HUMAN	P00414 homo sapien
38	66	46.5	261	1	COX3_PAPHA	O92xx8 papio hamad
39	66	46.5	261	1	COX3_RHIUN	O96065 rhinoceros
40	65	45.8	261	1	COX3_CARAU	O96133 carassius a
41	65	45.8	261	1	COX3_CYPCA	P15952 cyprinus ca
42	65	45.8	261	1	COX3_HALGR	P38597 halichoerus
43	65	45.8	261	1	COX3_HIPAM	O92425 hippopotamu
44	65	45.8	261	1	COX3_LATCH	O03170 latimeria c
45	65	45.8	261	1	COX3_PONPA	P92696 pongo pygma

ALIGNMENTS

RESULT 1
ID COX3_PARDE STANDARD; PRT; 273 AA.
AC P06030;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CYTOCHROME C OXIDASE POLYPEPTIDE III (EC 1.9.3.1) (CYTOCHROME AA3
SUBUNIT 3) (OXIDASE AA(3) SUBUNIT 3).
GN C1AE OR COIIT.
OS Paracoccus denitrificans.
OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC Paracoccus.
OX NCBI_TaxID=266;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-S1657;
RA Raitio M., Jalli T., Saraste M.;
RT "Isolation and analysis of the genes for cytochrome c oxidase in
Paracoccus denitrificans."
RL EMBO J. 6:2825-2833(1987).
RN [2]
RP SEQUENCE OF 1-143 FROM N.A.
RC STRAIN-S1657;
RX MEDLINE=87005242; PubMed=3019767;
RA Saraste M., Raitio M., Jalli T., Peraeama A.;
RT "A gene in Paracoccus for subunit III of cytochrome oxidase."
RL FEBS Lett. 206:154-156(1986).
RN [3]
RP SEQUENCE OF 1-15.
RX MEDLINE=88166732; PubMed=2832167;
RA Hailia T., Pustinen A., Finel M.;
RT "The Paracoccus denitrificans cytochrome a₃ has a third subunit."
RL Eur. J. Biochem. 172:543-546(1988).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
RX MEDLINE=95379947; PubMed=7651515;
RA Iwata S., Ostermeier C., Ludwig B., Michel H.;
RT "Structure at 2.8-A resolution of cytochrome c oxidase from
Paracoccus denitrificans."
RL Nature 376:660-669(1993).
CC -I- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
4 FERRICYTOCHROME C.
CC -I- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE.
CC -I- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC EMBL; X04406; CAA27995.1; -
CC EMBL; X05828; CAA29272.1; -


```

RX MEDLINE=85022618; PubMed=6091655;
RA Pepe G., Holtrop M., Gadaleta G., Kroon A.M., Cantatore P.,
RA Gallenzi R., de Benedetto C., Quagliariello C., Sbisa E.,
RA Saccocc C.;
RT "Non-random patterns of nucleotide substitutions and codon strategy
RT in the mammalian mitochondrial genes coding for identified and
RT unidentified reading frames";
RL Biochem. Int. 6:553-563(1983).
RN [3]
RP SEQUENCE FROM N.A.
RA STRAIN=SPRAGUE-DAWLEY; TISSUE=Liver;
RC Grosskopf R., Feldmann H.;
RT "Analysis of a DNA segment from rat liver mitochondria containing the
RT genes for the cytochrome oxidase subunits I, II, III, ATPase subunit
RT 6, and several tRNA genes.";
RL Curr. Genet. 4:151-158(1981).
RN [4]
RP SEQUENCE OF 1-10.
RC STRAIN=WISTAR; TISSUE=Liver;
RX MEDLINE=95324529; PubMed=7601105;
RA Schaeffer H., Noack H., Halangk W., Brandt U., von Jagow G.;
RT "Cytochrome c oxidase in developing rat heart. Enzymic properties and
RT amino-terminal sequences suggest identity of the fetal heart and the
RT adult liver isoform.";
RL Eur. J. Biochem. 230:235-241(1995).
CC -1- FUNCTION: SUBUNIT I, II, AND III FORM THE FUNCTIONAL CORE OF
CC THE ENZYME COMPLEX.
CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
CC 4 FERRICYTOCHROME C.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J01435; AAD15020.1; -
DR EMBL: M27315; AAB00994.1; -
DR EMBL: X14848; CAA32960.1; -
DR PIR: S04753; S04753.
DR HSSP: P00415; 10CC.
DR InterPro: IPR000298; Cytc_oxdse_III.
DR Pfam: PF00510; COX3; 1.
DR PROSITE: PS50253; COX3; 1.
KW Oxidoreductase; Mitochondrion; Transmembrane.
KM Oxidoreductase; Mitochondrion; Transmembrane.
RN INTL MET
FT CONFLICT 1 0 T -> S (IN REF. 4).
FT SEQUENCE 260 AA; 29739 MW; AAD4C8B2B777C3D3 CRC64;
SQ

```

Query Match 52.8%; Score 75; DB 1; Length 260;

Best Local Similarity 53.6%; Pred. No. 0.00021;

Matches 15; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 2 SIFLVCLIRIRLGAMSAKOHVFEMAA 29

DB 211 SIFLVCLIRIRLGAMSAKOHVFEMAA 238

```

RESULT 7
COX3_MOUSE
ID COX3_MOUSE STANDARD; PRT; 261 AA.
AC P00416;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CYTOCHROME C OXIDASE POLYPEPTIDE III (EC 1.9.3.1).
GN MTCO3 OR MT-CO3 OR COIII.
OS Mus musculus (Mouse).
OG Mitochondrion.

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=82137051; PubMed=7332926;
RA Bibb M.O., van Elten R.A., Wright C.T., Walberg M.W., Clayton D.A.;
RT "Sequence and gene organization of mouse mitochondrial DNA.";
RL Cell 26:167-180(1981).
CC -1- FUNCTION: SUBUNIT I, II, AND III FORM THE FUNCTIONAL CORE OF
CC THE ENZYME COMPLEX.
CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
CC 4 FERRICYTOCHROME C.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J01420; AAB48650.1; -
DR EMBL: V00711; CAA24090.1; ALT_TERM.
DR PIR: A00484; OTMS3.
DR HSSP: P00415; 10CC.
DR MGD: MGI:102502; mt-co3.
DR InterPro: IPR000298; mt-co3.
DR Pfam: PF00510; COX3; 1.
DR PROSITE: PS50253; COX3; 1.
KW Oxidoreductase; Mitochondrion; Transmembrane.
SQ SEQUENCE 261 AA; 29937 MW; DAF2EA445983D34E CRC64;

```

Query Match 52.8%; Score 75; DB 1; Length 261;

Best Local Similarity 53.6%; Pred. No. 0.00021;

Matches 15; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 2 SIFLVCLIRIRLGAMSAKOHVFEMAA 29

DB 212 SIFLVCLIRIRLGAMSAKOHVFEMAA 239

```

RESULT 8
COX3_PETMA
ID COX3_PETMA STANDARD; PRT; 261 AA.
AC O35539;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE CYTOCHROME C OXIDASE POLYPEPTIDE III (EC 1.9.3.1).
GN COXIII OR COIII.
OS Petromyzon marinus (Sea lamprey).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95229067; PubMed=7713438;
RA Lee W.J., Kocher T.D.;
RT "Complete sequence of a sea lamprey (Petromyzon marinus)
RT mitochondrial genome: early establishment of the vertebrate genome
RT organization.";
RL Genetics 139:873-887(1995).
CC -1- FUNCTION: SUBUNIT I, II, AND III FORM THE FUNCTIONAL CORE OF
CC THE ENZYME COMPLEX.
CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
CC 4 FERRICYTOCHROME C.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```


us-09-712-768-8.rsp

KW	Oxidoreductase; Mitochondrion; Transmembrane.
SQ	SEQUENCE 265 AA; 29328 MW; B10DI0921745180C CRC64;
Query Match	52.1%; Score 74; DB 1; Length 265;
Best Local Similarity	53.6%; Pred. No. 0.00031;

	Matches	15;	Conservative	3;	Mismatches	8;	Indels	0;	Gaps	0;
OY	2	SIFLAVCLIRILRGASAKOHVGFEMAA	29	:::: ::	::					
Dd	215	TFLPIVCGRQYLGHNLTKKNHYGFEAAA	242							

ID	COX3_ORYSA	STANDARD;	PRT;	265	AA.
AC	P14852;				
DT	01-APR-1990	(Rel. 14, Created)			

DE CYTOCHROME C OXIDASE POLYPEPTIDE III (EC 1.9.3.1).
GN COX3.

0C Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
0C Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
0C Ehrhartoideae; Oryzaeae; Oryza.

RP SEQUENCE FROM N.A.
RC STRAIN=CV, INDICA-IR36;
RX MEDLINE=90221829; PubMed=2158075;

RT subunit 3.¹,
RL Nucleic Acids Res. 18:371-371(1990).
RN [2]

RX MEDLINE=92315351; PubMed=1617739;
RA Liu A.W., Narayanan K.K., Andre C.P., Kaleikau E.K., Walbot V.;
RT "Co-transcription of orf25 and coxIII in rice mitochondria.";

```

CC      THE ENZYME COMPLEX,
CC      -I- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
CC      4 FERRICYTOCHROME C.

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use.

CC entitles requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
EMBL, Y17040, CAA234908.1.

DR PIR; JQ0166; QTRZM.
DR PIR; S20801; S20801.
DR PIR; S29748; S29748.
HSSD: P00415. 10CC

Oxidoreductase: Mitochondrion: Transmembrane
DR PROSTIN; P850253; COX3; 1.
DR Pfam; PF00510; COX3; 1.
DR InterPro IPR000298; CytC_oxdse_III.
DR

Query Match	52.18;	Score 74;	DB 1;	Length 265;
Best Local Similarity	53.68;			
Pred. No.	0.00031;			

Matches 15; Conservative 5; Mismatches 8; Indels 0; Gaps 0;
 Oy 2 SIFLVCLIRILRGAMSAKOHVFEEMAA 29
 Db 215 TLFVLICGIRGVLGHITKKHHVFEEMAA 242

RESULT 11
 COX3_ARBLI STANDARD; PRT; 109 AA.
 AC Q33752;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE CYTOCHROME C OXIDASE POLYPEPTIDE III (EC 1.9.3.1) (FRAGMENT).
 GN COIII.
 OS Arabacia lixula (Black urchin).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinzoa;
 OC Echinoidea; Euechinoidea; Echinacea; Arbacoidea; Arbacoidea; Arbacoidea.
 OX NCBI_TaxID=7640;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=92017217; PubMed=1681410;
 RA de Giorgi C., Lanave C., Musci M.D., Saccone C.;
 RT "Mitochondrial DNA in the sea urchin Arabacia lixula: evolutionary
 RT inferences from nucleotide sequence analysis.";
 RL Mol. Biol. Evol. 8:515-529(1991).
 CC -1- FUNCTION: SUBUNIT I, II, AND III FORM THE FUNCTIONAL CORE OF
 CC THE ENZYME COMPLEX.
 CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
 CC 4 FERRICYTOCHROME C.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M74839; AAA98044.1; -;
 DR InterPro: IPR000298; CytC_oxdse_III.
 DR Pfam: PF00510; COX3; 1.
 DR PROSITE: PS50253; COX3; 1.
 KW Oxidoreductase; Mitochondrion; Transmembrane.
 FT NON_TER 1
 SQ SEQUENCE 109 AA; 12511 MW; 3B4617A8F7512963 CRC64;

Query Match 51.4%; Score 73; DB 1; Length 109;
 Best Local Similarity 57.7%; Pred. No. 0.00019;
 Matches 15; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

Oy 4 FLVCLIRILRGAMSAKOHVFEEMAA 29
 Db 62 FLVCLIRILRGAMSAKOHVFEEMAA 87
 RESULT 12
 COX3_VICFA STANDARD; PRT; 265 AA.
 AC Q03227;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE CYTOCHROME C OXIDASE POLYPEPTIDE III (EC 1.9.3.1).
 GN COX3 OR COXIII.
 OS Vicia faba (Broad bean).
 OG Mitochondrion.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Vicia.
 OX NCBI_TaxID=3906;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90182277; PubMed=2155709;
 RA Macfarlane J.L., Wahlthner J.A., Wolstenholme D.R.;
 RT "A gene for cytochrome c oxidase subunit III (COXIII) in broad bean
 RT mitochondrial DNA: structural features and sequence evolution.";
 RL Curr. Genet. 17:33-40(1990).
 CC -1- FUNCTION: SUBUNIT I, II, AND III FORM THE FUNCTIONAL CORE OF
 CC THE ENZYME COMPLEX.
 CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
 CC 4 FERRICYTOCHROME C.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X51690; CAA35988.1; -;
 DR HSSP: P00415; IOCC.
 DR Mendel: 2198; VICFA:cox3;1.
 DR InterPro: IPR000298; CytC_oxdse_III.
 DR Pfam: PF00510; COX3; 1.
 DR PROSITE: PS50253; COX3; 1.
 KW Oxidoreductase; Mitochondrion; Transmembrane.
 SQ SEQUENCE 265 AA; 29468 MW; B3F0193DB8F0E9DD CRC64;

Query Match 51.4%; Score 73; DB 1; Length 265;
 Best Local Similarity 50.0%; Pred. No. 0.00044;
 Matches 14; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

Oy 2 SIFLVCLIRILRGAMSAKOHVFEEMAA 29
 Db 215 TLFVLICGIRGVLGHITKKHHVFEEMAA 242
 RESULT 13
 COX3_DROME STANDARD; PRT; 262 AA.
 AC P00417;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE CYTOCHROME C OXIDASE POLYPEPTIDE III (EC 1.9.3.1).
 GN MT:COIII OR COIII.
 GN Drosophila melanogaster (Fruit fly).
 OS Mitochondrion.
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE OF 1-179 FROM N.A.
 RA MEDLINE=83245048; PubMed=6408489;
 RA de Bruijn M.H.L.;
 RT "Drosophila melanogaster mitochondrial DNA, a novel organization and
 RT genetic code.";
 RL Nature 304:234-241(1983).
 RN [2]
 RP SEQUENCE OF 179-262 FROM N.A.
 RA MEDLINE=83220794; PubMed=6304652;
 RA Clary D.O., Wahlthner J.A., Wolstenholme D.R.;
 RT "Transfer RNA genes in Drosophila mitochondrial DNA: related 5'
 RT flanking sequences and comparisons to mammalian mitochondrial tRNA
 RT genes.";
 RL Nucleic Acids Res. 11:2411-2425(1983).
 RN [3]

RP SEQUENCE OF 179-262 FROM N.A.
 RC STRAIN-BRETAGNE;
 RX MEDLINE=88212147; PubMed=3130291;
 RA Garesse R.;
 RT "Drosophila melanogaster mitochondrial DNA: gene organization and
 evolutionary considerations.";
 RL Genetics 118:649-663(1988).
 CC -1- FUNCTION: SUBUNIT I, II, AND III FORM THE FUNCTIONAL CORE OF
 CC THE ENZYME COMPLEX.
 CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
 CC 4 FERROCYTOCHROME C.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: J01404; AAB59243.1; -
 DR EMBL: M37275; AAB69708.1; -
 DR EMBL: U37541; AAC47816.1; -
 DR PIR: B93477; OTFF3Y.
 DR PIR: S02250; S02250.
 DR HSSP: P00415; 10CC.
 DR FlyBase: FBgn0013676; mt:CoIII.
 DR InterPro: IPR000298; Cytc_oxdse_III.
 DR Pfam: PF00510; COX3; 1.
 DR PROSITE: PS50253; COX3; 1.
 KW Oxidoreductase; Mitochondrion; Transmembrane.
 SQ SEQUENCE 262 AA; 30064 MW; 68D5AEFC2297C130 CRC64;

Query Match 50.0%; Score 71; DB 1; Length 262;
 Best Local Similarity 57.7%; Pred. No. 0.00089;
 Matches 15; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

OY 4 FLVCLIRLRGMSAKOHVGFEMAA 29
 DB 215 FLVCLIRLRHNFHFKNHGFEMAA 240

RESULT 14
 COX3_DROYA STANDARD; PRT; 262 AA.
 AC P00418;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 15-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE CYTOCHROME C OXIDASE POLYPEPTIDE III (EC 1.9.3.1).
 GN COIII.
 OS Drosophila yakuba (Fruit fly).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7245;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=2317.6 IVORY COAST;
 RX MEDLINE=86089137; PubMed=3001325;
 RA Clary D.O.; Wolstenholme D.R.;
 RT "The mitochondrial DNA molecular of Drosophila yakuba: nucleotide
 sequence, gene organization, and genetic code.";
 RL J. Mol. Evol. 22:252-271(1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=83246544; PubMed=6306579;
 RA Clary D.O.; Wolstenholme D.R.;
 RT "Nucleotide sequence of a segment of Drosophila mitochondrial DNA
 that contains the genes for cytochrome c oxidase subunits II and III

RT and ATPase subunit 6.";
 RL Nucleic Acids Res. 11:4211-4227(1983).
 CC -1- FUNCTION: SUBUNIT I, II, AND III FORM THE FUNCTIONAL CORE OF
 CC THE ENZYME COMPLEX.
 CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
 CC 4 FERROCYTOCHROME C.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: X03240; CAA26990.1; -
 DR EMBL: X00924; CAA25443.1; -
 DR PIR: A00465; OTFF3.
 DR PIR: F25797; F25797.
 DR HSSP: P00415; 10CC.
 DR FlyBase: FBgn0013181; Dyak\mt:CoIII.
 DR InterPro: IPR000298; Cytc_oxdse_III.
 DR Pfam: PF00510; COX3; 1.
 DR PROSITE: PS50253; COX3; 1.
 KW Oxidoreductase; Mitochondrion; Transmembrane.
 SQ SEQUENCE 262 AA; 30039 MW; ED0C07F573666730 CRC64;

Query Match 50.0%; Score 71; DB 1; Length 262;
 Best Local Similarity 57.7%; Pred. No. 0.00089;
 Matches 15; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

OY 4 FLVCLIRLRGMSAKOHVGFEMAA 29
 DB 215 FLVCLIRLRHNFHFKNHGFEMAA 240

RESULT 15
 COX3_GADMO STANDARD; PRT; 261 AA.
 AC P55777;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE CYTOCHROME C OXIDASE POLYPEPTIDE III (EC 1.9.3.1).
 GN COXIII OR COIII.
 OS Gadus mohnua (Atlantic cod).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae;
 OC Gadus.
 OX NCBI_TaxID=8049;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NORWEGIAN COASTAL 1;
 RX MEDLINE=96414925; PubMed=8817926;
 RA Johansen S., Bakke I.;
 RT "The complete mitochondrial DNA sequence of Atlantic cod (Gadus
 mohnua): relevance to taxonomic studies among codfishes.";
 RL Mol. Mar. Biol. Biotechnol. 5:203-214(1996).
 CC -1- FUNCTION: SUBUNIT I, II, AND III FORM THE FUNCTIONAL CORE OF
 CC THE ENZYME COMPLEX.
 CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
 CC 4 FERROCYTOCHROME C.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC -----
DR EMBL; X99772; CAA68112.1; -
DR InterPro; IPR000298; CytC_oxdse_III.
DR Pfam; PF00510; COX3; 1.
DR PROSITE; PS50253; COX3; 1.
KW Oxidoreductase; Mitochondrion; Transmembrane.
SQ SEQUENCE 261 AA; 29719 MW; B6AA73F99E98BAF5 CRC64;

Query Match 48.6%; Score 69; DB 1; Length 261;
Best Local Similarity 50.0%; Pred. No. 0.0018;
Matches 14; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

OY 2 STELLVCLIRIRGAMSAKQHVGFEMAA 29
I I I I I I I I I I I I I I I I
Db 212 STEFLAVCLLRQIRYHFTSEHHGFEAAA 239

Search completed: January 10, 2002, 02:03:12
Job time: 204 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2002, 02:04:05 ; Search time 32.55 Seconds
(without alignments)
130.319 Million cell updates/sec

Title: US-09-712-768-8
Perfect score: 142
Sequence: 1 DSIFLVCLIRLTGMSAKOHGFEMMA 29

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues
Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SPREMBL_17.*
2: SP-archaea.*
3: SP-fungi.*
4: SP-human.*
5: SP-invertebrate.*
6: SP-mammal.*
7: SP-mhc.*
8: SP-organalle.*
9: SP-phage.*
10: SP-plant.*
11: SP-rodent.*
12: SP-virus.*
13: SP-vertebrate.*
14: SP-unclassified.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	87	61.3	290	2	Q9A304
2	86	60.6	298	2	Q9RM95
3	79	55.6	274	8	Q9G874
4	79	55.6	281	8	Q9G986
5	78	54.9	261	8	Q9Z231
6	77	54.2	272	8	Q9Z208
7	76	53.5	219	8	Q9S948
8	75	52.8	227	8	Q9B332
9	75	52.8	230	8	Q9B330
10	75	52.8	232	8	Q9B327
11	75	52.8	232	8	Q9B321
12	75	52.8	233	8	Q9G3G9
13	75	52.8	233	8	Q9G3G8
14	75	52.8	233	8	Q9G3G7
15	75	52.8	233	8	Q9G3G5
16	75	52.8	233	8	Q9G3F9
17	75	52.8	233	8	Q9G3F4
18	75	52.8	233	8	Q9G3F0
19	75	52.8	233	8	Q9G3E6

20	75	52.8	233	8	Q9G3E5	Q9G3E5	peromyscus
21	75	52.8	233	8	Q9G3E0	Q9G3E0	peromyscus
22	75	52.8	233	8	Q9G3D9	Q9G3D9	peromyscus
23	75	52.8	233	8	Q9G3D6	Q9G3D6	peromyscus
24	75	52.8	233	8	Q9G3D5	Q9G3D5	peromyscus
25	75	52.8	233	8	Q9G2H4	Q9G2H4	peromyscus
26	75	52.8	233	8	Q9G150	Q9G150	peromyscus
27	75	52.8	233	8	Q9G147	Q9G147	peromyscus
28	75	52.8	233	8	Q9G0V6	Q9G0V6	peromyscus
29	75	52.8	233	8	Q9G0V5	Q9G0V5	peromyscus
30	75	52.8	233	8	Q9G0V4	Q9G0V4	peromyscus
31	75	52.8	233	8	Q9B329	Q9B329	peromyscus
32	75	52.8	233	8	Q9B323	Q9B323	peromyscus
33	75	52.8	233	8	Q9B319	Q9B319	peromyscus
34	75	52.8	233	8	Q9B318	Q9B318	peromyscus
35	75	52.8	233	8	Q9B0S6	Q9B0S6	peromyscus
36	75	52.8	244	8	Q9M1V9	Q9M1V9	mesocricetu
37	75	52.8	261	8	Q9M0M4	Q9M0M4	mus musculu
38	75	52.8	261	8	Q9MDE6	Q9MDE6	mus musculu
39	75	52.8	262	8	Q9TBV9	Q9TBV9	actopora te
40	75	52.8	264	8	Q9T9H6	Q9T9H6	halocynthia
41	74	52.1	260	8	Q9MNMJ0	Q9MNMJ0	helobdella
42	74	52.1	261	8	Q9TA05	Q9TA05	lampectra fl
43	73	51.4	259	8	Q9MNMJ9	Q9MNMJ9	galathea lin
44	73	51.4	260	8	Q9G906	Q9G906	arabacia lix
45	73	51.4	261	8	Q9X131	Q9X131	vidua chaly

ALIGNMENTS

RESULT 1
ID Q9A304 PRELIMINARY; PRT; 290 AA.
AC Q9A304;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)
DE CYTOCHROME C OXIDASE, SUBUNIT III.
GN CC3402.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.
OX NCBI_Taxid=69394;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H., Berry K.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Fraser C.M.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.,
RT "Complete genome sequence of Caulobacter crescentus."
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL; AE006000; AAK25364.1; .
DR TIGR; CC3402; .
KW Complete proteome.
SQ SEQUENCE 290 AA; 32229 MW; 47016BD7090AB972 CRC64;

Query Match 61.3%; Score 87; DB 2; Length 290;
Best Local Similarity 53.6%; Pred. No. 3.5e-06;
Matches 15; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 2 SIFLVCLIRLTGMSAKOHGFEMMA 29
Db 237 TLFVLVCLIRLTGMSAKOHGFEMMA 264

RESULT 2
Q9RM95

RL Proc. R. Soc. Lond., B, Biol. Sci. 266:305-309(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Harild A.;
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: Y18522; CAA77201.1; -
 DR HSP: p18402; 1FFT.
 DR InterPro: IPR000298; CytC_oxdse_III.
 DR Pfam: PF00510; COX3; 1.
 DR PROSITE: PS50253; COX3; 1.
 KM Mitochondrion.
 FT NON_TER 261
 SQ SEQUENCE 261 AA; 29532 MW; 5021795FC7B5FD2 CRC64;

 Query Match 54.9%; Score 78; DB 8; Length 261;
 Best Local Similarity 53.6%; Pred. No. 9,4e-05;
 Matches 15; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

 QY 2 STELLVCLIRILRGAMSAKOHVFEMAA 29
 DB 212 STELLVCLIRILRGAMSAKOHVFEMAA 239

 RESULT 6
 092Z08 PRELIMINARY; PRT; 272 AA.
 AC 092Z08;
 DT 01-MAY-1999 (Tremblrel. 10, Created)
 DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE CYTOCHROME C OXIDASE POLYPEPTIDE III (EC 1.9.3.1).
 GN COX3.
 OS Cyanidioschyzon merolae.
 OC Mitochondrion.
 OC Eukaryota; Rhodophyta; Bangiophyceae; Cyanidioschyzon.
 OX NCBI_TaxID=45157;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=10D;
 RX MEDLINE=99030526; PubMed=9801318;
 RA Onita N., Sato N., Kuroiwa T.;
 RT "Structure and organization of the mitochondrial genome of the unicellular red alga Cyanidioschyzon merolae deduced from the complete nucleotide sequence."
 RT Nucleic Acids Res. 26:5190-5198(1998).
 RL EMBL: D89661; BAA34657.1; -
 DR HSP: p18402; 1FFT.
 DR InterPro: IPR000298; CytC_oxdse_III.
 DR Pfam: PF00510; COX3; 1.
 DR PROSITE: PS50253; COX3; 1.
 KM Oxidoreductase; Mitochondrion.
 FT NON_TER 272
 SQ SEQUENCE 272 AA; 31047 MW; ACEA7E88EF1F4B6C CRC64;

 Query Match 54.2%; Score 77; DB 8; Length 272;
 Best Local Similarity 50.0%; Pred. No. 0.00014;
 Matches 14; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

 QY 2 STELLVCLIRILRGAMSAKOHVFEMAA 29
 DB 222 STELLVCLIRILRGAMSAKOHVFEMAA 249

 RESULT 7
 095948 PRELIMINARY; PRT; 219 AA.
 AC 095948;
 DT 01-FEB-1997 (Tremblrel. 02, Created)
 DT 01-FEB-1997 (Tremblrel. 02, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE CYTOCHROME OXIDASE SUBUNIT III (FRAGMENT).
 GN COXIII.

OS Sepia elegans.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Sepioida;
 OC Sepiidae; Sepia.
 OX NCBI_TaxID=34527;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=97159560; PubMed=9007019;
 RA Bonnaud L., Boucher Rodoni R., Monnerot M.;
 RT "Phylogeny of cephalopods inferred from mitochondrial DNA sequences."
 RL Mol. Phylogenet. Evol. 7:44-54(1997).
 DR EMBL: Y07843; CAA69171.1; -
 DR InterPro: IPR000298; CytC_oxdse_III.
 DR Pfam: PF00510; COX3; 1.
 DR PROSITE: PS50253; COX3; 1.
 KM Mitochondrion.
 FT NON_TER 1
 SQ SEQUENCE 219 AA; 24743 MW; 2454BP6ECAD15E92 CRC64;

 Query Match 53.5%; Score 76; DB 8; Length 219;
 Best Local Similarity 57.1%; Pred. No. 0.00017;
 Matches 16; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

 QY 2 STELLVCLIRILRGAMSAKOHVFEMAA 29
 DB 192 STELLVCLIRILRGAMSAKOHVFEMAA 219

 RESULT 8
 09B332 PRELIMINARY; PRT; 227 AA.
 AC 09B332;
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE CYTOCHROME OXIDASE SUBUNIT III (FRAGMENT).
 GN COXIII.
 OS Peromyscus slevini (Slevin's mouse).
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Sigmodontinae;
 OC Peromyscus.
 OX NCBI_TaxID=44242;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=2439_CATALINA;
 RA Hainer D.V., Riddle B.R., Alvarez-Castaneda S.T.;
 RT "Evolutionary relationships of white-footed mice (Peromyscus) on islands in the Sea of Cortez, Mexico."
 RT Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 RL EMBL: AF343755; AAK32167.1; -
 KM Mitochondrion.
 FT NON_TER 1
 SQ SEQUENCE 227 AA; 25524 MW; E2157F384CD68029 CRC64;

 Query Match 52.8%; Score 75; DB 8; Length 227;
 Best Local Similarity 53.6%; Pred. No. 0.00025;
 Matches 15; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

 QY 2 STELLVCLIRILRGAMSAKOHVFEMAA 29
 DB 200 STELLVCLIRILRGAMSAKOHVFEMAA 227

 RESULT 9
 09B330 PRELIMINARY; PRT; 230 AA.
 AC 09B330;
 DT 01-JUN-2001 (Tremblrel. 17, Created)

DT 01-JUN-2001 (Tremblrel. 17, last sequence update)
 DE CYTOCHROME OXIDASE SUBUNIT III (FRAGMENT).
 GN COIII.
 OS Peromyscus interparietalis.
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Sigmodontinae;
 OC Peromyscus.
 NCBI_TaxID=152342;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=2438_SALSTIPUES;
 RA Hafner D.J., Riddle B.R., Alvarez-Castaneda S.T.;
 RT "Evolutionary relationships of white-footed mice (Peromyscus) on
 islands in the Sea of Cortez, Mexico."
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF343757; AAK32169.1; -
 KW Mitochondrion.
 FT NON_TER 1 1
 FT 230 230
 SQ SEQUENCE 230 AA; 25894 MW; DPAACDCEFE6571A CRC64;

Query Match 52.8%; Score 75; DB 8; Length 230;
 Best Local Similarity 53.6%; Pred. No. 0.00026;
 Matches 15; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 2 STELLVCLIRILRGAMSAKOHVGFEMAA.29
 DB 203 STELLVCLIRILRGAMSAKOHVGFEMAA.230

RESULT 10
 O9B327 PRELIMINARY; PRT; 232 AA.
 AC O9B327;
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, last sequence update)
 DE CYTOCHROME OXIDASE SUBUNIT III (FRAGMENT).
 GN COIII.
 OS Peromyscus eremicus (cactus mouse).
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Sigmodontinae;
 OC Peromyscus.
 NCBI_TaxID=42410;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=2489_TIBURON;
 RA Hafner D.J., Riddle B.R., Alvarez-Castaneda S.T.;
 RT "Evolutionary relationships of white-footed mice (Peromyscus) on
 islands in the Sea of Cortez, Mexico."
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF343760; AAK32172.1; -
 KW Mitochondrion.
 FT NON_TER 1 1
 FT 232 232
 SQ SEQUENCE 232 AA; 26064 MW; FCIJD518CAC4E6489 CRC64;

Query Match 52.8%; Score 75; DB 8; Length 232;
 Best Local Similarity 53.6%; Pred. No. 0.00026;
 Matches 15; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 2 STELLVCLIRILRGAMSAKOHVGFEMAA.29
 DB 205 STELLVCLIRILRGAMSAKOHVGFEMAA.232

RESULT 11
 O9B321

ID O9B321 PRELIMINARY; PRT; 232 AA.
 AC O9B321;
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, last sequence update)
 DE CYTOCHROME OXIDASE SUBUNIT III (FRAGMENT).
 GN COIII.
 OS Peromyscus collatus.
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Sigmodontinae;
 OC Peromyscus.
 NCBI_TaxID=152343;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=2437_DAVITL;
 RA Hafner D.J., Riddle B.R., Alvarez-Castaneda S.T.;
 RT "Evolutionary relationships of white-footed mice (Peromyscus) on
 islands in the Sea of Cortez, Mexico."
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF343768; AAK32180.1; -
 KW Mitochondrion.
 FT NON_TER 1 1
 FT 232 232
 SQ SEQUENCE 232 AA; 26191 MW; A0491D58DAA3A813 CRC64;

Query Match 52.8%; Score 75; DB 8; Length 232;
 Best Local Similarity 53.6%; Pred. No. 0.00026;
 Matches 15; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 2 STELLVCLIRILRGAMSAKOHVGFEMAA.29
 DB 205 STELLVCLIRILRGAMSAKOHVGFEMAA.232

RESULT 12
 O9G3G9 PRELIMINARY; PRT; 233 AA.
 AC O9G3G9;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-JUN-2001 (Tremblrel. 17, last sequence update)
 DE CYTOCHROME OXIDASE SUBUNIT III (FRAGMENT).
 GN COIII.
 OS Peromyscus boylii (brush mouse).
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Sigmodontinae;
 OC Peromyscus.
 NCBI_TaxID=56316;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC PubMed=11083930;
 RA Riddle B.R., Hafner D.J., Alexander L.F.;
 RT "Phylogeography and Systematics of the Peromyscus eremicus Species
 Group and the Historical Biogeography of North American Warm Regional
 Deserts."
 RL Mol. Phylogenet. Evol. 17:145-160(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC PubMed=11095731;
 RA Riddle B.R., Hafner D.J., Alexander L.F., Jaeger J.R.;
 RT "Cryptic vicariance in the historical assembly of a Baja California
 Peninsular Desert biota."
 RL Proc. Natl. Acad. Sci. U.S.A. 97:14438-14443(2000).
 DR EMBL: AY009175; AAG45506.1; -
 DR InterPro: IPR000298; CYC_oxdase_III.
 DR Pfam: PF00510; COX3; 1.
 DR PROSITE: PS50253; COX3; 1.
 KW Mitochondrion.
 FT NON_TER 1 1
 FT 233 233

SO SEQUENCE 233 AA; 26243 MW; 8C728AD659F18954 CRC64;

Query Match 52.8%; Score 75; DB 8; Length 233;
Best Local Similarity 53.6%; Pred. No. 0.00026;
Matches 15; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

OY 2 STFLVCLIRILRGAMSAKOHVFEMAA 29
| | | | | : : : | | | | |
DB 206 STFLVCLIRQLKHFHTSKHHFGFEAAA 233

RESULT 13

OY9G3G8 PRELIMINARY; PRT; 233 AA.

AC OY9G3G8: 01-MAR-2001 (TREMBlrel. 16, Created)
RT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DR CYTOCHROME OXIDASE SUBUNIT III (FRAGMENT).
GN COIII.
OS Peromyscus californicus (California mouse).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Sigmodontinae;
OC Peromyscus.
OX NCBI_TaxID=42520;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=11083930;
RA Riddle B.R., Hafner D.J., Alexander L.F., Jaeger J.R.;
RT "Phylogeography and Systematics of the Peromyscus eremicus Species
Group and the Historical Biogeography of North American Warm Regional
Deserts.";
RL Mol. Phylogenet. Evol. 17:145-160(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX PubMed=11095731;
RA Riddle B.R., Hafner D.J., Alexander L.F., Jaeger J.R.;
RT "Cryptic vicariance in the historical assembly of a Baja California
Peninsular Desert biota.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:14438-14443(2000).
DR EMBL: AY009176; AAG45507.1; -
DR InterPro: IPR000298; Cytc_oxdse_III.
DR Pfam: PF00510; COX3; 1.
DR PROSITE: PS50253; COX3; 1.
GN CYTOCHROME OXIDASE SUBUNIT III (FRAGMENT).
GN COIII.
OS Peromyscus merriami.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Sigmodontinae;
OC Peromyscus.
OX NCBI_TaxID=144753;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=11083930;
RA Riddle B.R., Hafner D.J., Alexander L.F., Jaeger J.R.;
RT "Cryptic vicariance in the historical assembly of a Baja California
Peninsular Desert biota.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:14438-14443(2000).
DR EMBL: AY009177; AAG45508.1; -
DR InterPro: IPR000298; Cytc_oxdse_III.
DR Pfam: PF00510; COX3; 1.
DR PROSITE: PS50253; COX3; 1.
GN Mitochondrion.
FT NON_TER 1 1
FT SEQUENCE 233 AA; 26233 MW; A370E9DCFD8F6831 CRC64;

Query Match 52.8%; Score 75; DB 8; Length 233;
Best Local Similarity 53.6%; Pred. No. 0.00026;
Matches 15; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

OY 2 STFLVCLIRILRGAMSAKOHVFEMAA 29
| | | | | : : : | | | | |
DB 206 STFLVCLIRQLKHFHTSKHHFGFEAAA 233

RESULT 15

OY9G3G5 PRELIMINARY; PRT; 233 AA.

AC OY9G3G5: 01-MAR-2001 (TREMBlrel. 16, Created)
RT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DR CYTOCHROME OXIDASE SUBUNIT III (FRAGMENT).
GN COIII.
OS Peromyscus merriami.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Sigmodontinae;
OC Peromyscus.
OX NCBI_TaxID=144754;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=11083930;
RA Riddle B.R., Hafner D.J., Alexander L.F., Jaeger J.R.;
RT "Phylogeography and Systematics of the Peromyscus eremicus Species
Group and the Historical Biogeography of North American Warm Regional
Deserts.";
RL Mol. Phylogenet. Evol. 17:145-160(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX PubMed=11095731;
RA Riddle B.R., Hafner D.J., Alexander L.F., Jaeger J.R.;
RT "Cryptic vicariance in the historical assembly of a Baja California
Peninsular Desert biota.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:14438-14443(2000).
DR EMBL: AY009179; AAG45510.1; -
DR InterPro: IPR000298; Cytc_oxdse_III.
DR Pfam: PF00510; COX3; 1.
DR PROSITE: PS50253; COX3; 1.
GN Mitochondrion.
FT NON_TER 1 1
FT SEQUENCE 233 AA; 26233 MW; 2EA98BA643813A34 CRC64;

Query Match 52.8%; Score 75; DB 8; Length 233;
Best Local Similarity 53.6%; Pred. No. 0.00026;
Matches 15; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

OY 2 STFLVCLIRILRGAMSAKOHVFEMAA 29
| | | | | : : : | | | | |
DB 206 STFLVCLIRQLKHFHTSKHHFGFEAAA 233

RESULT 14

OY9G3G7 PRELIMINARY; PRT; 233 AA.

AC OY9G3G7: 01-MAR-2001 (TREMBlrel. 16, Created)
RT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DR CYTOCHROME OXIDASE SUBUNIT III (FRAGMENT).
GN COIII.
OS Peromyscus crinitus.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

	TITLE	JOURNAL	COMMENT
RESULT 4 BE600048/c LOCUS	BE600048	529 bp mRNA,	EST 18-AUG-2000
DEFINITION	PtL78_GG06.bJ_A002 Pathogen induced 1 (PtI) Sorghum bicolor cDNA,	mRNA sequence.	
ACCESSION	BE600048		
VERSION	BE600048.1	GI:9855121	
KEYWORDS	EST.		
SOURCE	sorghum.		
ORGANISM	Sorghum bicolor		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Sorghum. 1 (bases 1 to 529)		
AUTHORS	Cordonnier-P Pratt,M.-M., Gingle,A., Dean,R., Sudman,M. and Pratt .L.H.		
	An EST database from Sorghum: pathogen-induced plants Unpublished (2000) Contact: Cordonnier-Pratt MM Department of Botany The University of Georgia Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA Tel.: 706 542 1860 Fax: 706 542 1805 Email: mmpratt@uga.edu		
	Sequences have been trimmed to exclude polyA, vector and regions below phred quality 16. The threshold for highest quality sequences is 20.		
	Seq primer: JEN REV High quality sequence stop: 468 polyA-No.		

```

FEATURES
source
Location/Qualifiers
1..529
/organism="Sorghum bicolor"
/db_xref="taxon:4558"
/clone_id="Pathogen induced 1 (P11)"
/note="Organ: Anthracnose-infected leaves from
two-week-old sorghum plants 48 hr after inoculation;
Vector: pBluescript II from lambda zap II; Site.1: XhoI;
Site.2: EcoRI. Two-week-old sorghum plants (BTx 623
cultivar) were infected with pathogen (isolate FR421 of
Colletotrichum graminicola, which is a sorghum isolate).
RNA was prepared from infected leaves harvested from 45
seedlings (2 weeks old) exhibit juvenile resistant
reaction, which is an incompatible interaction. As they
grow older (4 weeks or older), plants resume susceptibility
to anthracnose disease. The library was made from poly-A
RNA in the cloning vector lambda ZAP II. Clones to be
sequenced were prepared by mass excision. WARNING: While
most or all ESTs are expected to derive from the host
plant, no effort was made to eliminate ESTs deriving from
the pathogen."
71 a 204 c 146 g 108 t

```

Query Match	34.5%	Score 30	DB 10	Length 529
Best Local Similarity	61.5%	Pred. No. 22		
Matches 48	Conservative 0	Mismatches 30	Indels 0	Gaps 0
Qy	8	tcttcctgcgtgcctcctcgcgcgcacgcgcgcgcgtgcgagtgcgcgaacacagcacg	67	
Db	288	ttcctcctcttccatctgcgtcgtcttctgctcctcctcgcgcgcgtccatcccggtaaagacggcgcg	229	
Qy	68	tcgatttcagagatgcgcg	85	
Db	228	ggcggccgagacgatccgcc	211	

RESULT	5			
BF152072				
LOCUS	BF152072	271 bp	mRNA	EST
				29-DEC-2000

DEFINITION	u224c07.y1 NCI-CGAP Mam5 Mus musculus cDNA clone IMAGE:366996 5' similar to SM-CAL3 MOUSE P08121 COLLAGEN ALPHA 1(I)II CHAIN
ACCESSION	BF152072.1
VERSION	GI:11033467
KEYWORDS	EST.
SOURCE	house mouse.
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL	Unpublished (1997)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: Lochar Hennighausen Ph.D., Robin Humphreys cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNUd) DNA sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNUd at: image.llnl.gov/image/html/tiresources.shtml
FEATURES	MGI:1430764 Trace considered overall poor quality Seq primer: ~40RP from Gibco High quality sequence stop: 1. Location/Qualifiers 1..271
SOURCE	

	/organism="Mus musculus"
	/strain="C57/B6"
	/db_xref="taxon:10090"
	/clone="IMAGE:366996"
	/clone_1ib="NCI_CGAP_Mam5"
	/tissue_type="Tumor, gross tissue"
	/dev_stage="7 months"
	/lab_host="DH10B"
	/note="Orig.: mammary; Vector: PCW-SPOrt6; site_1: SalI;
	site_2: NotI; Cloned unidirectionally. Primer: Oligo dt.
	library constructed by Life Technologies. Investigators
	providing samples: Lothar Hennighausen/Robin Humphreys,
	NH."
BASE COUNT	34 a 77 c 80 g 80 t
ORIGIN	
Query Match	34.3%; Score 29.8; DB 11; Length 271;
Best Local Similarity	63.0% Pred. No. 22;
Matches 46; Conservative	0; Mismatches 27; Indels 0; Gaps 0;
Dy	11 ttccgcgtgctgccatcgacatcccggcggaggcatgttcggcaaacagcacgtcg 70
Db	67 TCCGGCTGCCTCCCCGTGTCATAGTGTTGATGGTAGTGATCGTGGACGCCCTGGTGCTCC 126
OY	71 gtttcgagatgac 83
Db	127 TCATTCTGCTGGC 139

RESULT		6	
AA728883/c			
LOCUS			
DEFINITION	337 bp	mRNA	EST
	nv37d11.l	NCL-CCGAP-B5 Homo sapiens cDNA clone IMAGE:122389	22-JUN-1998
	s similar to SW:P2C2_ARATH P49598 PROTEIN PHOSPHATASE 2C ; , mRNA		
	sequence.		
ACCESSION	AA728883		
VERSION	AA728883.1	GI:2750242	
KEYWORDS	EST .		
SOURCE	human.		
ORGANISM	Homo sapiens		
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			


```
/organism="Homo sapiens"  
/db_xref="taxon:9606"  
/clone="IMAGE:780950"  
/clone_1lb="Soares_testis_NHT"  
/sex="male"
```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:725807"
/clone_lib="Scares ovary tumor NBH01"
/sex="Female"
/tissue_type="ovarian tumor"
/lab_host="DH10B (ampicillin resistant)"
/node="Organ: ovary: Vector: pRT73D (Pharmacia) with a modified polylinker: Site.1: Not I; Site.2: Eco RI; Site.3: Xba I"

```


REFERENCE AUTHORS TITLE JOURNAL COMMENT	SOURCE
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium. 1 (bases 1 to 663) Carlton, J.M.-R. and Dame, J.B. The Plasmodium vivax and P. berghei gene sequence tag projects Parasitol. Today 16 (10), 409 (2000) Contact: Dame JB Dept. of Pathobiology, College of Veterinary Medicine University of Florida 2015 SW 23rd Avenue, Bldg 1017, Gainesville, FL 32611, USA Tel.: 352 392 4700 Fax: 352 392 9704 Email: damej@mail.ufl.edu Seq primer: M13(-20) forward Class: Shotgun.	Location/Qualifiers 1. .663 /organism="Plasmodium vivax" /strain="Salvador I (Collins, W. 1972, J. Parasitol. 69, 497-598)" /db_xref="taxon:5885" /clone_id="Pv MRN #30"

/dev_stage="asexual blood forms"
 /lab_host="Salmirid bolivensis"
 /note="Vector: pBluescript SK(+) vector DNA, phagemid
 excised from lambda ZAP; Site_1: EcoR V; Site_2: EcoR V;
 Host leukocytes were extracted from P. vivax infected
 blood using the following methods: first, infected blood
 was activated by the addition of 0.5 ml of ADP (40ng/ml)
 per 10 ml blood. Then blood was passed over a column of
 acid washed 0.1 mm glass beads, then through a Plasmid
 filter, followed by passage through a column of pre-wet
 Whatman CFI powder (1:2 ratio volume of blood to CFI),
 and finally centrifuged through a 50% Percoll density
 cushion. Purified DNA was digested with mung bean nucleases
 in the presence of 44% formamide at 500C as described
 (Venick, K.D., Imberski, R.B., and McCutchan, T.F. 1988.
 Nucleic Acids Research 16: 6883-6896). Digested DNA was
 blunt-ended using T4 DNA polymerase and size fractionated
 over a Sepharose CL-2B column. Fractions in the size range
 500bp-4kb were ligated into the Eco RV site of pBluescript
 SK(+), and E. coli XL-10 Gold transformed with the
 ligation mixture."

BASE COUNT 111 a 192 c 168 g 192 t
 ORIGIN

Query Match 32.9%; Score 28.6; DB 13; Length 663;
 Best Local Similarity 61.3%; Prid. No. 61;
 Matches 46; Conservative 0; Indels 0; Cuts 0

Accession	Version	Keywords	Organism	LOCUS	Definition	Result 15
U00001	1	plasmid	Escherichia coli	U00001	U00001	U00001
U00002	1	plasmid	Escherichia coli	U00002	U00002	U00002
U00003	1	plasmid	Escherichia coli	U00003	U00003	U00003
U00004	1	plasmid	Escherichia coli	U00004	U00004	U00004
U00005	1	plasmid	Escherichia coli	U00005	U00005	U00005
U00006	1	plasmid	Escherichia coli	U00006	U00006	U00006
U00007	1	plasmid	Escherichia coli	U00007	U00007	U00007
U00008	1	plasmid	Escherichia coli	U00008	U00008	U00008
U00009	1	plasmid	Escherichia coli	U00009	U00009	U00009
U00010	1	plasmid	Escherichia coli	U00010	U00010	U00010
U00011	1	plasmid	Escherichia coli	U00011	U00011	U00011
U00012	1	plasmid	Escherichia coli	U00012	U00012	U00012
U00013	1	plasmid	Escherichia coli	U00013	U00013	U00013
U00014	1	plasmid	Escherichia coli	U00014	U00014	U00014
U00015	1	plasmid	Escherichia coli	U00015	U00015	U00015
U00016	1	plasmid	Escherichia coli	U00016	U00016	U00016
U00017	1	plasmid	Escherichia coli	U00017	U00017	U00017
U00018	1	plasmid	Escherichia coli	U00018	U00018	U00018
U00019	1	plasmid	Escherichia coli	U00019	U00019	U00019
U00020	1	plasmid	Escherichia coli	U00020	U00020	U00020
U00021	1	plasmid	Escherichia coli	U00021	U00021	U00021
U00022	1	plasmid	Escherichia coli	U00022	U00022	U00022
U00023	1	plasmid	Escherichia coli	U00023	U00023	U00023
U00024	1	plasmid	Escherichia coli	U00024	U00024	U00024
U00025	1	plasmid	Escherichia coli	U00025	U00025	U00025
U00026	1	plasmid	Escherichia coli	U00026	U00026	U00026
U00027	1	plasmid	Escherichia coli	U00027	U00027	U00027
U00028	1	plasmid	Escherichia coli	U00028	U00028	U00028
U00029	1	plasmid	Escherichia coli	U00029	U00029	U00029
U00030	1	plasmid	Escherichia coli	U00030	U00030	U00030
U00031	1	plasmid	Escherichia coli	U00031	U00031	U00031
U00032	1	plasmid	Escherichia coli	U00032	U00032	U00032
U00033	1	plasmid	Escherichia coli	U00033	U00033	U00033
U00034	1	plasmid	Escherichia coli	U00034	U00034	U00034
U00035	1	plasmid	Escherichia coli	U00035	U00035	U00035
U00036	1	plasmid	Escherichia coli	U00036	U00036	U00036
U00037	1	plasmid	Escherichia coli	U00037	U00037	U00037
U00038	1	plasmid	Escherichia coli	U00038	U00038	U00038
U00039	1	plasmid	Escherichia coli	U00039	U00039	U00039
U00040	1	plasmid	Escherichia coli	U00040	U00040	U00040
U00041	1	plasmid	Escherichia coli	U00041	U00041	U00041
U00042	1	plasmid	Escherichia coli	U00042	U00042	U00042
U00043	1	plasmid	Escherichia coli	U00043	U00043	U00043
U00044	1	plasmid	Escherichia coli	U00044	U00044	U00044
U00045	1	plasmid	Escherichia coli	U00045	U00045	U00045
U00046	1	plasmid	Escherichia coli	U00046	U00046	U00046
U00047	1	plasmid	Escherichia coli	U00047	U00047	U00047
U00048	1	plasmid	Escherichia coli	U00048	U00048	U00048
U00049	1	plasmid	Escherichia coli	U00049	U00049	U00049

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 10, 2002, 00:56:35 ; Search time 1374.03 Seconds

(without alignments)
1044.558 Million cell updates/sec

Title: US-09-712-768-7

Perfect score: 87
Sequence: 1 gattcgatctctctggtglt.....tcggttcgagatgcgcga 87

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 segs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenDbml:*

- 1: gb_ba:*
- 2: gb_hlg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vl:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_om:*
- 20: em_or:*
- 21: em_ov:*
- 22: em_pat:*
- 23: em_ph:*
- 24: em_pl:*
- 25: em_ro:*
- 26: em_sts:*
- 27: em_sy:*
- 28: em_un:*
- 29: em_vl:*
- 30: em_hgo_hum:*
- 31: em_hgo_inv:*
- 32: em_hgo_rod:*
- 33: em_hgo_hum:*
- 34: em_hgo_inv:*
- 35: em_hgo_rod:*
- 36: em_hgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	87	100.0	87	6	AX113992	AX113992 Sequence
2	48.4	55.6	286550	1	SME591785	AL591785 Sinorhizo
3	46.8	53.8	4890	1	PDCOX1	X05828 Paracoccus
4	45.4	52.2	10835	1	AE008010	AE008010 Agrobacte
5	43.6	50.1	12299	1	AE006000	AE006000 Caulobact
6	43.6	50.1	346510	1	AP003011	AP003011 Mesorhizo
7	42	48.3	8121	1	BUA242592	AJ242592 Bradyrhiz
8	42	48.3	31495	1	BU033883	U33883 Bradyrhizob
9	35.2	40.5	12829	1	AE004449	AE004449 Pseudomon
10	31	35.6	294800	1	SME591789	AL591789 Sinorhizo
11	30.4	34.9	227	8	AY023455	AY023455 Oryza sat
12	30.2	34.7	714	9	MACCY03A	M85148 Macaca mula
13	30	34.5	17683	9	AC019055	AC019055 Homo sapi
14	29.8	34.3	125110	9	AL450267	AL450267 Human DNA
15	29.8	34.3	185147	2	AP002388	AP002388 Homo sapi
16	29.2	33.6	83463	9	AL161644	AL161644 Human DNA
17	29.2	33.6	151552	2	AL357673	AL357673 Homo sapi
18	28.6	32.9	16472	9	HMTTCSEQ	X92256 Hylobates l
19	28.4	32.6	1325	1	MB062766	U62766 Mycobacteri
20	28.4	32.6	110000	2	LMFCHR36_02	Continuation (3 of
21	28.4	32.6	110000	2	LMFCHR36_03	Continuation (4 of
22	28.4	32.6	209175	2	AL450321	AL450321 Mus muscu
23	28.2	32.4	5642	1	AF228578	AF228578 Rhizobium
24	28	32.2	690	10	AY010181	AY010181 Amnosperr
25	28	32.2	690	10	AY010184	AY010184 Amnosperr
26	28	32.2	690	10	AY010185	AY010185 Amnosperr
27	28	32.2	690	10	AY010186	AY010186 Amnosperr
28	28	32.2	690	10	AY010194	AY010194 Amnosperr
29	27.8	32.0	2147	5	AB005533	AB005533 Colurnix
30	27.6	31.7	1591	3	AF397532	AF397532 Drosophill
31	27.6	31.7	1604	3	AF218237	AF218237 Drosophill
32	27.6	31.7	9429	2	AC020441	AC020441 Drosophill
33	27.6	31.7	16884	4	TEUY19192	Y19192 Talpa europ
34	27.6	31.7	37750	1	SC6D7	AL133213 Streptomy
35	27.6	31.7	110000	2	LMFCHR28_0	AL390935 Leishmani
36	27.6	31.7	110000	2	LMFCHR32_05	Continuation (6 of
37	27.6	31.7	156508	3	AC007581	AC007581 Drosophill
38	27.6	31.7	170089	3	AC007925	AC007925 Drosophill
39	27.6	31.7	189370	1	AF010496	AF010496 Rhodobact
40	27.6	31.7	194815	9	CNS01DYZ	AL136295 Human chr
41	27.6	31.7	328500	3	AE003465	AE003465 Drosophill
42	27.4	31.5	2814	1	AF315510	AF315510 Lysobacte
43	27.4	31.5	27644	3	CEIK0282	U41558 Caenorhaddi
44	27	31.0	12363	4	AE002088	AE002088 Deinococc
45	27	31.0	16561	9	PTWITG	X93335 Pan troglod

ALIGNMENTS

RESULT 1
AX113992 87 bp DNA
LOCUS AX113992
DEFINITION Sequence 7 from Patent EP1103603.
ACCESSION AX113992
VERSION AX113992.1 GI:13940152
KEYWORDS
SOURCE
ORGANISM
Glucanobacter oxydans.
Glucanobacter oxydans
Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
Glucanobacter.

REFERENCE
1 (bases 1 to 87)

AUTHORS Asakura, A., Hoshino, T. and Shinjoh, M.

Cytochrome c oxidase complex from glucanobacter oxydans

JOURNAL Patent: EP 1103603-A 7 30-MAY-2001;

F. HOFMANN-IA ROCHE AG (CH)

FEATURES

location/Qualifiers
1..87
/organism="Glucanobacter oxydans"
/db_xref="taxon:442"


```

/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="60 KD CHAPERONIN A PROTEIN"
/protein_id="CAC45364.1"
/db_xref="GI:15073723"
/translation="MAAEKVEFGSAREKMLRGVYLADAVKVTILGPRGNVVDKSF
GAPRTTKDGVSAKEIELEDEKFNENGAQMREVASKTNIDAGDGTATVLAQAIHRE
GAKAAAGMNPMDLKGIDILAAEVYKDLAKRKINTSDENAVGTTISANGEQITGL
DIAEMOKVMEGVITVEAKTAELEEVYGMKPDGYSIPFTVTPKMWADLEDA
FILLHEKLSNQLQMLPYLEAVVOTGKFLTIADVGEALATLVNKLKGLKIAV
KAPGDRKRMALDIALITGTVISBDLGIKLSVTLDMIGRAKKSITKENTIID
GAGKSDIEGRVAQIKQIETTSIDYDREKLOERLAKLAGVAIVRGATGEVVK
KDRIDALNATRAVQEGIVGCGVALIRSSVKITVKGENDODAGVNIYRALQSPA
ROIYVNAQDEASIVYKRIEKNDDPGYNQOTGSGIMANGIIDPYKAVRTALQDA
SVASLTITTEMIMELPKKDA PAMPGMGMGKGDMA"
complement(4298..4594)
/gene="groES1 OR SMC00912"
complement(4298..4594)
/gene="groES1 OR SMC00912"
/function="cell processes: chaperoning"
/note="predicted by Codon_usage
predicted by Homology
predicted by Framed"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="10 KD CHAPERONIN A PROTEIN"
/protein_id="CAC45365.1"
/db_xref="GI:15073724"
/translation="MASTNRPPLHDRVVRVSESEKTKGIIIPDTAKKPOEGEIV
AVSGAGDESGKVPIDVKAGDRILFGKMSGTEVKINGEDLLIMKEADINGVIG"
4863..5690
/gene="SMC00911"
4863..5690
/gene="SMC00911"
/function="miscellaneous: hypothetical/global homology"
/note="Product confidence: hypothetical
Gene name confidence: hypothetical
predicted by Codon_usage
predicted by Framed"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="CONSERVED HYPOTHETICAL PROTEIN"
/protein_id="CAC45366.1"
/db_xref="GI:15073725"
/translation="MASFNSISSEKLARLVGTRKAPVIVDIPDDDPADHILVPSF
HRGHVHPQAEYRGWVYVLCRHGKLSGSAAMLRHGAAGAEILDGIEMRAG
LPLVPASLPNSDPAGSVWVTRSRPKIDRIACPMILRRVDPARFLVTPAEVAV
GERRGATPFDIEGVFWSHRECECTFDTMVEGLSPALPHVARIYRGATDREREIP
EAGGLAVSLIGLSRMYADDLEOLEAGMTIYDALYRVARADATQEKHDVSHSGKGRK"
6004..6852
/gene="SMC00910"
6004..6852
/gene="SMC00910"
/function="miscellaneous: hypothetical/global homology"
/note="Product confidence: hypothetical
Gene name confidence: hypothetical
predicted by Codon_usage
predicted by Homology
predicted by Framed"
/codon_start=1
/transl_table=11
/evidence=not_experimental
/product="CONSERVED HYPOTHETICAL PROTEIN"
/protein_id="CAC45367.1"
/db_xref="GI:15073726"
/translation="MAVRINSEREIAGRYDVVLCVGVILHNGIQAFASACALAEAR
AGGLTVLITNSRPHGVYVQIRGLGVDEAVDRIVTSGDVQALIAADRIFPIG
ADRDLPLEGLTEIYCADEAETIVCAFYDDETEPEHYRAITLGLARKIPIICAN
PDLVVERGHRILPCAGAIKAIYELGEGARIGKPKYIATYRAALAEAKARAPFQISR
VIAIGDMPTDVAGADAGFDLLIYISAGIHAQEIYHESRTDARIMAFILRQNGAQPKW

```

```

gene      MPRLA"
6866..7849
/gene="ribf OR SMC00909"
6866..7849
/gene="ribf OR SMC00909"
/EC_number="2.7.1.26"
/EC_number="2.7.1.2"

Query Match      55.6%; Score 48.4; DB 1; Length 286550;
Best Local Similarity 74.4%; Pred. No. 0.00036;
Matches 61; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Db 130642 CGATCTTCCTCTGCTGCTGCTGCGCGCCCTCGGCGGCGACTTCACCGCAAGCACCC 130701
QY 65 acgtcggttcagatgagcgccg 86
|| ||| ||||| || |||
Db 130702 ACTTCGGCTTCGAGCGCGCGCC 130723

RESULT 3
LOCUS      PDCOX1      4890 bp      DNA      BCT      12-SEP-1993
DEFINITION Paracoccus denitrificans cytochrome c oxidase CoII and CoIII genes
coding for major subunits II and III.
ACCESSION X05828
VERSION X05828.1 GI:45468
KEYWORDS  CoII gene; cytochrome aaz; cytochrome c oxidase; unidentified
reading frame.
SOURCE      Paracoccus denitrificans.
ORGANISM    Paracoccus denitrificans.
Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
Paracoccus.
REFERENCE 1 (bases 1 to 4887)
AUTHORS    Raitio,M., Jallil,T. and Saraste,M.
TITLE      Isolation and of the genes for cytochrome c oxidase in Paracoccus
denitrificans
JOURNAL    EMBO J. 6, 2825-2833 (1987)
REFERENCE 2 (bases 1 to 4890)
AUTHORS    Saraste,M.
TITLE      Direct Submission
JOURNAL    Submitted (07-MAR-1988) to the EMBL/Genbank/DBJ databases
COMMENT    Data kindly reviewed (07-MAR-1988) by Saraste M.
FEATURES
source
1..4890
/organism="Paracoccus denitrificans"
/db_xref="taxon:266"
/map="cox1 locus"
853..856
/note="gccc was ggc in [1]"
/citation=[1]
862..863
/note="ag was a in [1]"
/citation=[1]
880..882
/note="agt was at in [1]"
/citation=[1]
892..894
/note="tgg was tg in [1]"
/citation=[1]
898..900
/note="gcc was gc in [1]"
/citation=[1]
900..906
/note="inverted repeat A"
912..915
/note="TATA-box like sequence"
924
/note="g was c in [1]"
/citation=[1]
937..938
/note="tg was tcgc in [1]"

```

```

repeat_unit      940..946      /citation=[1]
                  /note="Inverted repeat A'"
RBS              984..987      /note="pot. ribosome binding site"
                  992..1888      /note="coiled gene product (AA 1 - 298)"
CDS              /codon_start=1
                  /transl_table=11
                  /protein_id="CAA29268.1"
                  /db_xref="GI:45469"
                  /translation="MAIATKRGVAAVWSLGVATMTAVPALADVDLPVIGKRVN
GMNRPASSPLAHQOMIDHPVLYITPAVTIFVCLLLICIVRRNRANPAPRPTH
NPTETVITLVPLVLIIVAGASPLIFRSQMPNDPDVIAIGHQRTSTERYNDG
VAEDALMLKEALADAGYSEDEYLLATDPVVPVPGKRVLYVGTATDVTHATITPFA
VKODAPGRILADLWFSVDEGVYFGGCSLGCINHAHYMPEIVKAVSOEKYEMLAGAK
EEPAADASYLPAPVKLASAE"
RBS              1961..1965
                  /note="pot. ribosome binding site"
CDS              1974..2813
                  /note="ORF 1 (AA 1 - 279)"
                  /codon_start=1
                  /transl_table=11
                  /protein_id="CAA29269.1"
                  /db_xref="GI:45470"
                  /translation="MSLVFTAFVGLMIAPQVNPVPAFCVAFLTALGGASGALNMW
YDADIDAVMRKTRGPRVPSGIVTSOEPPLAVGIALSGLSYMLGAGNMFAFLFTI
FFAVAVYTTIMLKRTSPQNIIVIGAGAPPMIGALPTGIGIESILMFALIFETTP
HEWALAFMKDDSKAGVPMILVTYHGRKVTCHIFAYTIVLAPFALMLGFTSVGPLY
LAVSVLNLFLTAGWQILRSEDOQAQAGYVEKRYFLSLYTFPLHLTALVQHWY
GGN"
RBS              2802..2806
                  /note="pot. ribosome binding site"
CDS              2813..3082
                  /note="ORF 2 (AA 1 - 89)"
                  /codon_start=1
                  /transl_table=11
                  /protein_id="CAA29270.1"
                  /db_xref="GI:45471"
                  /translation="MLPVEHLELHKRRSRNIGLVLLAFVALFGLSVKRTGDM
MGGYDHRPRASMLRPDPDPAPAPAAVAPGPAANQVGTGEANQ"
CDS              3079..3663
                  /note="ORF 3 (AA 1 - 194)"
                  /codon_start=1
                  /transl_table=11
                  /protein_id="CAA29271.1"
                  /db_xref="GI:45472"
                  /translation="MSGGKRSTRTVYVMAAGVYVLMGALSNAAVPEYSPCKVNGFA
GTTNVAEASDVYDERIRYRFDANSDSLGTFRRPQREMLKIGENALTYEALIN
TDEPVTGTASVYVNPDAAGYFNKIECFCTEQTLDPGEVEKPSFEVDADLVNDRD
AVIRIDITLSTYFRTDPPAPKQAALDAKTEPVN"
RBS              3686..3691
                  /note="pot. ribosome binding site"
CDS              3698..4522
                  /note="coiled gene product (AA 1 - 274)"
                  /codon_start=1
                  /transl_table=11
                  /protein_id="CAA29272.1"
                  /db_xref="GI:45473"
                  /translation="MAHVKNHDYQILLPSPITWPFGAIGAVMLTGAVAMKGTTFEGL
PVBGPWMEILIGLVLYVMFGWADVVBSETEHTRPVYRIGIQYFIIFINSEVWF
VAMFAMRIKALYPMGPDSPIDKGVMPREGIVTEWHMLPLINTLILLSGVAVTWAF
HAFVLEBGRKTTINGLIVAILGVCFTGQAYEYSHAAGLADTVYAGAFYVATGPHG
AHVIGTIFLFCVLLIRLKGOMTOKOHVGEAAMAWMHVYDVVWMLFLFVYITMGR"
                  4543..4549
                  /note="inverted repeat B"
                  4543..4574
                  /note="pot. transcription terminator"

```

```

repeat_unit      4559..4565
                  /note="inverted repeat B'"
BASE COUNT      777 a 1581 c 1623 g 909 t
ORIGIN
Query Match      53.8%; Score 46.8; DB 1; Length 4890;
Best Local Similarity 73.2%; Pred. No. 0.002;
Matches 60; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
QY              5 cgaattctcgtgctgctgcgtatccgcatcctgcgctgcgattcgcgcaaacagc 64
                  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB              4371 CGATCTCTCTGTCCTGCTGATCCGCTCAAGGCCAGATGACCCGAGAGAGC 4430
                  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY              65 acgtcgattcgaagatgcccgc 86
                  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB              4431 ATGTCGGCTTCGAGGCCGCCGC 4452
                  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT          4
AE008010        10835 bp      DNA      BCT      14-AUG-2001
LOCUS           Agrobacterium tumefaciens strain C58 circular chromosome, section
DEFINITION      68 of 254 of the complete sequence.
ACCESSION       AE008010 AE007869
VERSION         AE008010.1 GI:15155740
KEYWORDS        Agrobacterium tumefaciens.
SOURCE          Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
ORGANISM        Rhizobiaceae; Rhizobium.
REFERENCE       1 (bases 1 to 10835)
AUTHORS         Hinkle, G., Slater, S.C. and Goodner, B.
TITLE           Complete Genome Sequence of Agrobacterium tumefaciens C58
                (Rhizobium radiobacter C58), the Causative Agent of Crown Gall
                Disease in Plants
JOURNAL         Unpublished
REFERENCE       2 (bases 1 to 10835)
AUTHORS         Hinkle, G., Slater, S.C. and Goodner, B.
TITLE           Direct Submission
JOURNAL         Submitted (14-AUG-2001) Cereon Genomics, 45 Sidney Street,
                Cambridge, MA 02139, USA
FEATURES        Location/Qualifiers
                1..10835
                /organism="Agrobacterium tumefaciens"
                /strain="C58"
                /db_xref="taxon:358"
                /complement(118..642)
                /gene="AGR_C_1396"
                /complement(118..642)
                /gene="AGR_C_1396"
                /note="(AJ242592) hypothetical protein"
                /codon_start=1
                /transl_table=11
                /product="AGR_C_1396p"
                /protein_id="AAK86575.1"
                /db_xref="GI:15155741"
                /translation="MRISPLARTFLAAGLAIIPAAALAAQQPSTPKSTHGSVY
                ICDKPGASBEOCALMONVIADREVELSVYILKTADRSRIILAPLGLKDM
                ELYIDNNNIGRAYFTGRCSBECYVEVDIDDLVLVLRGKNAVALRESVDQDVGLP
                IELSGFARGYDALP"
                988..1938
                /gene="AGR_C_1397"
                /note="(AJ242592) cytochrome c oxidase subunit II"
                /codon_start=1
                /transl_table=11
                /product="AGR_C_1397p"
                /protein_id="AAK86576.1"
                /db_xref="GI:15155742"
                /translation="MDSASVIDLDQTRREGGVNRIYALAGMTCLLTAVGAADDP
                VHWQMGQPAATPIMHEIRWFEQYLTWIVYVTLFVALLIIVAKFAAKNPVASKT

```

```

gene
CDS
/sequence="SHNTAIEVWTLLAVLILLFLAPSPFNLLNALTOPENPDLLKATATONLSYEYKA
AEGAEPLSEDSYLLTKDODRAAEKEDKARYPRILLAYDNEMVVPYKTVRLLYTAAFD
VIAHAFAMPARGVKIDAVPGRLENTWPRPEKRGILYGCSEICGKDHAFMLAIRVSE
QOYTWTHAAAASDINGANRALMASVDGAPRTVDVAAENETN"
2549. .3664
/gene="AGR_C_1399"
2549. .3664
/gene="AGR_C_1399"
/sequence="CYTOCHROME C OXIDASE POLYPEPTIDE I (CYTOCHROME AA3
SUBUNIT 1)"
/codon_start=1
/transl_table=11
/product="AGR_C_1399p"
/protein_id="AAK86577.1"
/db_xref="GI:15155743"
/translation="MDAIFSLHVAAGASSILGAINFTTTLLNKRAGMTLHKMPLERAW
SVLYTAPALLSLPVLAGGITMLLTNDNFTAPFSPGGDPILYOLHFWFPGHPEYV
ILLPGFGLISHIVSTFSKRPVGLGMAYAMAIGAVGIYMAHMYITVGLSLAQR
YFVATVAVIAPVPGIKIJSWIAIWMGSLFSTPMWVAIGFIFLETYGVGYOLANA
GLDSRLDHYVVAHFHYVLGLAVFAIFAGWYWEPKITGYWNEFISGLHFMVMTI
GVNLIFPPHGLIAGMPRRYIDPDYAGMMVSSYGSYISAVAVGIFLFGWEAFA
KRIAGNPMGCEGATILEMQLSPPTTHOEDLPRT"
4054. .5007
/gene="AGR_C_1402"
4054. .5007
/gene="AGR_C_1402"
/sequence="(X89566) heme O synthase"
/note="(X89566) heme O synthase"
/codon_start=1
/transl_table=11
/product="AGR_C_1402p"
/protein_id="AAK86578.1"
/db_xref="GI:15155744"
/translation="MTVIDRDMWGAASSFLSPAGADYPELLKPRVMSLYVTARAG
LVLAGGINPILGILITICITAGAGASGAINMYDADIDVMTAKRPLPSRIAPR
EALAFGLTASPSVITGLVNNFSGLLFTTFYAVVTMLKSTPONTIIGAA
GAPRPLGMACVGVGSLSDSVILFLITLTPAHFALALFKMDGAVGAIPIPMWVA
GERSTKNQMTVAVLTAANAAPFETGLSAGYIPAAVLSAIFVYCSFVRRMPGSD
EKMLPAKMFAYSVLYLFAIFGLADHFAPLAKVIAISGL"
5167. .5784
/gene="AGR_C_1406"
5167. .5784
/gene="AGR_C_1406"
/sequence="AGR_C_1406"
/note="CYTOCHROME C OXIDASE ASSEMBLY PROTEIN CTAG"
/codon_start=1
/transl_table=11
/product="AGR_C_1406p"
/protein_id="AAK86579.1"
/db_xref="GI:15155745"
/translation="MANMEAEETARAGSORVNSRSLVLCVFEPCAMIGMAVAAVPLYS
LFCRVYNGTTOREYSDVILDKTINVFADNTSGLMMDRPVDKMKPKIGETI
KVFKAATNRSPVATGTAAVNVTPMEAGAVFNVEGCEFTTLOPGBTLEMPVFEV
DPDIATARETKSHVTLISTYTPYPAKTEKPVASLPKTESGSKL"
5846. .6721
/gene="AGR_C_1407"
5846. .6721
/gene="AGR_C_1407"
/sequence="AGR_C_1407"
/note="(A0242592) cytochrome c oxidase subunit III"
/codon_start=1
/transl_table=11
/product="AGR_C_1407p"
/protein_id="AAK86580.1"
/db_xref="GI:15155746"
/translation="MADTHOKNHDYHIIIDSPMPLASIGAFIMFGGVCYVARYLSSG
SFLFGAEELANPMLFYLIVLYVYMAWMDTIGKANGSHRYVSLHRLRGIMHFI
ASVFMFVAMFAYPDAISLPHPAIGASLEUTGGWPKGEVYIDPMHLPYLNVTIL
LLSGTCVTNAHNAHLNDRKGLISGLATVALGVLESTVOVEYIHAPEDFKNSIYGA
TFPMATGFHFHVFGLVCLFRAIAGGFTPKOHGFEEAAMWYMHVDVWMLFLF
FALYIWGNGAPLHG"
6941. .7369
/gene="AGR_C_1409"
6941. .7369
/gene="AGR_C_1409"
/sequence="AGR_C_1409"
/note="(A0242592) hypothetical protein"

```

```

gene
CDS
/sequence="
/codon_start=1
/transl_table=11
/product="AGR_C_1409p"
/protein_id="AAK86581.1"
/db_xref="GI:15155747"
/translation="MYSRGCRCREREVLMSPEPSNGNFAPVYDIKYLKCCPCRCNGR
LPDGLTPKPCACAGLDYGFADAGDPAVFWLLVGLVVGMLMVDORFAPVMAH
WMMLPFYIVISVTLRLKLGIMIALOYRNNAESEGLDRE"
7338. .8114
/gene="AGR_C_1411"
7338. .8114
/gene="AGR_C_1411"
/sequence="(AF182952) SufEit 1"
/note="(AF182952) SufEit 1"
/codon_start=1
/transl_table=11
/product="AGR_C_1411p"
/protein_id="AAK86582.1"
/db_xref="GI:15155748"
/translation="MQAKGLIYNDVSHKKAORVWFAPIVLLALILGLGTWQV
KRIYKEALMADIERRNSPATLSDIEATVSGEIEYRKVRLSGTDFHREHEFA
THQGGYGYITVPLTLADRLILVNRGFEVPEPMKAARPDGVSGETYITGLAAPL
VAKPSSLDPNDIANKNIEFYKDLAAMASSAEVPEPLVKLFVDADNAPNPGMPCGV
TLIDLPNNHLOXAITWYGLAALVIAGRAYFRNCKAGE"
8173. .9219
/gene="AGR_C_1414"
8173. .9219
/gene="AGR_C_1414"
/sequence="(AE006203) LytB"
/note="(AE006203) LytB"
/codon_start=1
/transl_table=11
/product="AGR_C_1414p"
/protein_id="AAK86583.1"
/db_xref="GI:15155749"
/translation="MNIAVSKPEPLTIRLCPGKFCAGYDRAIQIYVILAKYAPVY
RHEIYHNRVYVGLBAKGAIEVEEHEIIPAEHREQVVEFSAHVKSVEPAOANLF
YLDACPLSKVAKQAMRORIGRNVILIGHAGREPVITGMLGPBGTSVIVVEYDA
GYVEPRDNEFVOTQITSDVDTAGVIAAROEPRPAIOAPDAISICATITRQDAYK
QAARCDLEIYVAGARNSSKRLVYVALRAQKSHVLYORASEIDMNTGIDRTVGLS
AGASAPVYVDEIIEAFKARPDITDLAVVEETHEFLVNLRSIELTTDDMAFVNG
NASNALPVPKATAGI"
9312. .10280
/gene="AGR_C_1416"
9312. .10280
/gene="AGR_C_1416"
/sequence="AGR_C_1416"
/note="HOMOSERINE KINASE (HK)"
/codon_start=1
/transl_table=11
/product="AGR_C_1416p"
/protein_id="AAK86584.1"
/db_xref="GI:15155750"
/translation="MAVYTDITEDELNPLFOYDVGSULTSYKGLAEGVENSFLIHTT
KQPLITTYEKREKNDLPFLGLMOMHLAAGLSQPLPKRDELLOGLSGPRALI
SFLBGMWMLRKPRAKCHREVGAALAAHMLASGEFELIKRNALSYDGMKYLMDSEERAD
EVEKGLREIPEIDYLAAMHPKRDPAVIAHADLEODMVFELGSLIDIFYACND
LLAYDVSIQIANMCPKEDGAYVWTKGKALLGYSYVRLSRAELBALPILSRGSAIAR
FLTRLYDMTLTPAGALVYKRPDELYLRKLRHRTIANAEYGLAGE"
10277. .10717
/gene="AGR_C_1417"
10277. .10717
/gene="AGR_C_1417"
/sequence="R1bonuclease H (E.C.3.1.26.4) Mutant with His 62
Replaced by Pro (h62p)"
/note="R1bonuclease H (E.C.3.1.26.4) Mutant with His 62
Replaced by Pro (h62p)"
/codon_start=1
/transl_table=11
/product="AGR_C_1417p"
/protein_id="AAK86585.1"
/db_xref="GI:15155751"
/translation="MKHYDITPDGACSGNPGPGGAVIARYGETEKELSGCEATITNN
RMELIATISALMAKSPCEVDLYTDSATVKGDTIKWIFGMAKKGWKTADNKPVANVEL
WQALEAQRERKRVTLHWYKGAGHGPENRADELARKGMEPKRR"
3194 a 3242 g 2378 t

```

BASE COUNT

2021 a

Query Match	52.2%	Score 45.4;	DB 1;	Length 10835;
Best Local Similarity	73.4%;	Pred. No. 0.0044;		
Matches	58;	Conservative	0;	Mismatches 21; Indels 0; Gaps 0;

QY	8	tcttcctcgtcgtctgcatacccgatccgcatccctcgcggtgatgtcggcaaacagcacg	67
Db	6552	TCTTCTCCTCGTGTCGCTGTGGCCGCCGCACATCGCGCGGTTTCACACCACAAGCATATT	6611
OY	68	tcggttccgaatgagcgacg	86
Db	6612	TGGGTTTCGAAGCGGCCCC	6630

RESULT	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
5	AEO06000/c	Caulobacter crescentus section 326 of 359 of the complete genome.	AEO06000	AE005673	AE006000.1 GI:13425102	Caulobacter crescentus. Caulobacter crescentus. Bacteria: Proteobacteria; alpha subdivision; Caulobacter group; Caulobacter	28-MAR-2001
REFERENCE	AUTHORS						
1 (bases 1 to 12299)	Nierman,W.C., Feldblum,T.V., Laub,M.T., Paulsen,I.T., Nelson,K.E., Eisen,J., Heidelberg,J.F., Alley,M.R.K., Ohta,N., Maddock,J.R., Potocka,I., Nelson,W.C., Newton,A., Stephens,C., Phadke,N.D., Ely,B., DeBoy,R.T., Dodson,R.J., Durkin,A.S., Gwinn,M.L., Haft,D.H., Kolonay,J.F., Smit,J., Craven,M., Khouli,H., Shetty,J., Berry,K., Uterback,T., Tran,K., Wolf,A., Vamathevan,J., Ermolaeva,M., White,O., Salzberg,S.L., Venter,J.C., Shapiro,L. and Fraser,C.M.						
JOURNAL	Proc. Natl. Acad. Sci. U.S.A.	98 (7), 4136-4141	(2001)				
MEDLINE	2 (bases 1 to 12299)	Nierman,W.C., Feldblum,T.V., Paulsen,I.T., Nelson,K.E., Eisen,J., Heidelberg,J.F., Alley,M.R.K., Ohta,N., Maddock,J.R., Potocka,I., Nelson,W.C., Newton,A., Stephens,C., Phadke,N.D., Ely,B., Laub,M.T., DeBoy,R.T., Dodson,R.J., Durkin,A.S., Gwinn,M.L., Haft,D.H., Kolonay,J.F., Smit,J., Craven,M., Khouli,H., Shetty,J., Berry,K., Uterback,T., Tran,K., Wolf,A., Vamathevan,J., Ermolaeva,M., White,O., Salzberg,S.L., Shapiro,L., Venter,J.C. and Fraser,C.M.					
REFERENCE	AUTHORS						
TITLE	Complete genome sequence of Caulobacter crescentus						
JOURNAL	Proc. Natl. Acad. Sci. U.S.A.	98 (7), 4136-4141	(2001)				
MEDLINE	2 (bases 1 to 12299)	Nierman,W.C., Feldblum,T.V., Paulsen,I.T., Nelson,K.E., Eisen,J., Heidelberg,J.F., Alley,M.R.K., Ohta,N., Maddock,J.R., Potocka,I., Nelson,W.C., Newton,A., Stephens,C., Phadke,N.D., Ely,B., Laub,M.T., DeBoy,R.T., Dodson,R.J., Durkin,A.S., Gwinn,M.L., Haft,D.H., Kolonay,J.F., Smit,J., Craven,M., Khouli,H., Shetty,J., Berry,K., Uterback,T., Tran,K., Wolf,A., Vamathevan,J., Ermolaeva,M., White,O., Salzberg,S.L., Shapiro,L., Venter,J.C. and Fraser,C.M.					
REFERENCE	AUTHORS						
TITLE	Direct Submission						
JOURNAL	Submitted (31-JAN-2001) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA						
FEATURES	source						
gene	organism="Caulobacter crescentus"						
CDS	/db_xref="taxon:69394" 102..1913 /gene="CC3393" 102..1913 /note="Identified by match to PFAM protein family HMM PF00441" /codon_start=1 /transl_table=1 /product="acyl-CoA dehydrogenase family protein" /protein_id="AAK2535.1" /db_xref="GI:13425103"						
translation="MATTGAPLKGARGPSPMTFRFVRDLASLRHVAFERLADPPEADDAVAALVEAGAFASDIILAPIRRODVLGARLENSVPRAPGFADAYQOFAGKWTSLAPBAHGSGGLPKTELEVAVLEWQAANNAFGICPMISGAILPALBCSPBGKAKTLGVSEWGTMNLTPEPGASDIADLAATTPGDGGMRTIGQKITMGDDH AADNVHLVARLPDPAPRGVGISIFLPAKVLDGDTGIPANMLRASLSHKGIHGSPTCVMIPEGAKELVGLGOGIPMFPMMAARLVNGTOGVTAIERAYOQALAFSOCSRAQGRSANTGAYPSRLPDHPRYRRITLYLMKHIAEARSICSTAYAADLARAADPAARTAAALREQLIPTAKAWSTDIGWVASOGVOIHGGSGFIETEGAQHTROARKPLAAYECTGRIQALIDIVGRKLIMEGQAVGDLMDIDIRTDITDLKASDVSGTRLEBPADAAT							

gene
CDS
complement(1984..2466)
/gene="CC3394"
complement(1984..2466)
/gene="CC3394"
/note="identified by match to PFAM protein family HMM PF00578"

gene
CDS
2595..3182
/gene="CC3395"
2595..3182
/gene="CC3395"
/note="identified by glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAK25357.1"
/db_xref="GI:13425105"
/translation="MASLIDDDGGEFLGRLLVAMPGITDPRFERYLYLCADHEDANNA
MGLANRPEVEGLATPRLDALDMAASAIKRPRSLLAGYAGWGPOLDELRDVMYLCDPA
DEGLIFDEDHDKMTALAKLGITADHLSTAGRA"
complement(3250..4917)
/gene="CC3396"
complement(3250..4917)
/gene="CC3396"
/note="identified by match to protein family HMM"
/codon_start=1
/transl_table=11
/product="GGDF family protein"
/protein_id="AAK25358.1"
/db_xref="GI:13425106"
/translation="MSFTGERRIWDATATLEAGADVALLMEPETDRILNGAAR
ALGIDPLPECCSAFRALALPODRAQAEVLKPREPGSEVAVARRVGGETCLMRGEG
MLEEVRAGAAYVAETPKFSASELCDLTGLDRRSFLARERLAOEGTHQLVVLDLPB
LRRLNLEIGHERRADLVLAIGSGRLAAPAOISIIGRIGEEFEVALCOPIGYEPDVLL
SALDEPLVACGDHPHTLSIGAVSREGGLDAPDAELLRRALEAVEANAARGGAAMA
YGRAMETGSLALEADLRGAIGKETTPYQPYRVIRSTALSFEELAWIIPRRGS
MLPREFPLLEEGLMSBELGAHHMHAAQULSTRANHPPAGNLIVSNLSTEIDB
POLVDVATEVTLVNRNLRPGALKLEVESDIMDPERRAVILKTURLDAGAGIALDFGTG
GSSSYLYRLPFDTLKIDRYEVRTMGNNASAKIYRSVVKLGDLDLEVAEDEVENNA
EMAHLOSIGCDYGOGFGPARPALSPQEAEVYLNEAYVYDGAAPVKARG"
complement(4952..5530)
/gene="CC3397"
complement(4952..5530)
/gene="CC3397"
/note="identified by match to protein family HMM"
/codon_start=1
/transl_table=11
/product="ribosomal-protein-alanine acetyltransferase,
putative"
/protein_id="AAK25359.1"
/db_xref="GI:13425107"
/translation="MSVLIPLSPRRRLLOGRAVYLLRPQRDYAPWAALSRSSRAI
LBPMPTEWHEDLSRAAFRAALAARELECEAVPTFPREDALLGALRHVRHRRAL
GSLTIGTYWLGEPYRGVGMADAVENTLIRPAFIGLGHRLLEAACMPENHASALLL
KGFSEEGTAHYXIKINGAMRDHRLEGLVAPE"
complement(5548..6819)
/gene="CC3398"
complement(5548..6819)
/gene="CC3398"
/note="identified by match to PFAM protein family HMM
PF00675"
/codon_start=1


```

Db 341331 CCATCTTCTGCTGCTTCTGATCCGCGCATGAAGCGCATTTTACCCCAAGCAGC 341390
Oy 65 acgtcggttcgaagatgcccgc 86
Db 341391 ATTTCGGCTTCGAGCGCGCCG 341412

RESULT 7
BJA242592 8121 bp DNA BCT 30-SEP-1999
LOCUS Bradyrhizobium japonicum coxB, coxA, coxE, coxF, coxG, coxH, sbh1
DEFINITION genes, partial t1d gene and ORFs 133 and 177.
ACCESSION AJ242592
VERSION AJ242592.1 GI:6006409
KEYWORDS assembly protein; coxA gene; coxB gene; coxC gene; coxE gene; coxF
gene; coxG gene; cytochrome c oxidase subunit I; cytochrome c
oxidase subunit II; cytochrome c oxidase subunit III; heme A
synthase; Heme O synthase; hypothetical protein; ORF133; protome
IX farnesyltransferase; sbh1 gene; SUR1 homolog; t1d gene; t1d
homolog.

SOURCE Bradyrhizobium japonicum.
ORGANISM Bradyrhizobium japonicum.
Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
Bradyrhizobium group; Bradyrhizobium.
REFERENCE 1 (bases 1 to 8121)
AUTHORS Rossmann, R., Iofere, H., Rossi, P. and Hennecke, H.
TITLE Factors involved in biogenesis of active cytochrome a3 encoded by
the coxBPFC gene cluster from Bradyrhizobium japonicum
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 8121)
AUTHORS Rossmann, R.
TITLE Direct Submission
JOURNAL Submitted (20-MAY-1999) Rossmann R., Eidgenossische Technische
Hochschule Zuerich, Mikrobiologisches Institut, Schmelzbergstrasse
7, CH-8092 Zuerich, SWITZERLAND
Location/Qualifiers
FEATURES
source 1. 8121
/organism="Bradyrhizobium japonicum"
/strain="110spc4"
/db_xref="taxon:375"
gene complement(1..440)
CDS /gene="t1d"
complement(<1..440)
/gene="t1d"
/function="putative role in regulation of carbon storage"
/transl_table=1
/codon_start=1
/product="t1d homolog"
/protein_id="CAB56816.1"
/db_xref="GI:6006410"
/translation="MTNPDYTSLLDRANLDRDVRHREVAAGLAGADGELFLFYSQTE
ALMFNGRLKQATYDTSQSGRLRAVDAVGAHSDVSLPALIRADVAVAVRGYS
GSFSAPEPHNPLVGDNDPLAPGETVYKLAIEDATLRKDP"
complement(542..1075)
/feature="ORF177"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="CAB56817.1"
/db_xref="GI:6006411"
/translation="MGLSRKMARPPSSSLNALIGLAAALLALPSLAHNOGAVRSVHG
DMOIRCDYPPGAOACALIQSVAVEDRSNAGLITVILKTADOKSLMLVAPLGLVL
PSGLIKDNODVGRAGFVRCPLPNCVAVVMDKLLGOLRTAKATATFIIFETPEGI
GEPPLSLNGIGEGYDKLP"
1523..2362
/gene="coxB"
1523..1612
/gene="coxB"
1523..2362
/gene="coxB"
1523..2362
/gene="coxB"
/EC_number="1.9.3.1"
/function="cytochrome c oxidation (respiration)"

/codon_start=1
/transl_table=1
/evidence="experimental"
/product="cytochrome c oxidase subunit II"
/protein_id="CAB56818.1"
/db_xref="GI:6006412"
/translation="MRSSKRGVRGRIHLLGLAVGLTATGSAAPAEIGCPAPMWTLOQ
SGSPVMDNIWVFNHFLFVLIATITFVLALVLYVITKFNARANPVPSRTNHTLLEVA
WTLVPVLIIVGISVPSFRLIFELDPKADITIKATKOWMSYAPDPNGKEPDSLM
AODKOPRLGVDNEMVVPYKVIKRVQVGTADVIAPALPAGVKIDALPGRNEMWPK
AAKTGMFYGCSELCKDHAFMFIARVVEDEMFASWVETAKKKRSGGTGYASAG
PNO"
2451..4076
/gene="coxA"
2451..4076
/gene="coxA"
/EC_number="1.9.3.1"
/function="cytochrome c oxidation (respiration)"
/codon_start=1
/transl_table=1
/evidence="experimental"
/product="cytochrome c oxidase subunit I"
/protein_id="CAB56819.1"
/db_xref="GI:6006413"
/translation="MATAAHGDHNAQDHDEHAHPGWRVYVSTNHKDIGTYLI
FVAVIGVGAAMSIAIRAEIATPVQVIFHEHTYVAVTSHGLMIFPMVPMATGFR
GMFVPLMGAPDMAFPPRMNLSFWLLPASEGLMSTFVEBPGANGVAGMTYVP
LSSSGHPVAVPFAISHLAAGASSILGATITITFMNAPAGMTLHKMPLFWSIIV
IVFLLISLPLVLAGATMLTPDRNGTFEAPDGGDVLHQLHFMFGHPREVLTIL
PQGMISOIVTSFESRSPVRCYLGMAVYNAITGICFPVYMAHMTYVGSATQAVFA
ATMVIAVPTGVKIFSNIAITMGSGSTIEFRAPKIMAVGTFILFTVGSVTVLANAGDR
VLOETTYVAHFHYVLSGAVAFAGWYVPMKMTGYMETLAKAHFWTFGLVNL
VEFPQHFGLISGMPRRYVDPDAFAGMNVSSVSGVYISGFLVILFYCVIDAFAKKVP
ACDNPWGAATLEWTLPSPPHFQFVLPVQ"
4137..5081
/gene="coxE"
4137..5081
/gene="coxE"
/function="protoheme IX farnesyltransferase"
/codon_start=1
/transl_table=1
/product="putative heme O synthase"
/protein_id="CAB56820.1"
/db_xref="GI:6006414"
/translation="MSVLDQNAVNDINPISAEVGDYIALLKPRYMSLVITFALVGA
NAPGHFHPVLAITSLCIAVAGASGALNMLEGIDDKMSRTANRPPIRGRITRPEA
MTFGMTLAFESVMTLGIIVNWIAGALAFITFVYVITMMLKMTAONITVIGAGGA
LPPVYMAAVTGVDEPPLLEPAIFETPPHMAALFRSDDVARAGIIPMLPYNAP
DATROIILTYTIVLIAVANAAPMAIGPFAVYGVSLIAGAMLVLAIVNYRRRSOS
LRAPKTRFASITVIFALPATLLAEVYRVALAPMAGCA"
5078..5251
/gene="coxF"
5078..5251
/gene="coxF"
/function="cytochrome c oxidase"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="CAB56821.1"
/db_xref="GI:6006415"
/translation="MIAVADKPEPDGIVLTEAOKSRRSRIAIALGLVLYLFEAV
TMVKGPAVLVRPM"
5257..5901
/gene="coxG"
5257..5901
/gene="coxG"
/function="biogenesis of cytochrome c oxidase"
/codon_start=1
/transl_table=1
/product="experimental"
/evidence="experimental"
/product="putative heme A synthase or assembly protein"
/protein_id="CAB56822.1"
/db_xref="GI:6006416"
/translation="MDHEPTISRDSRTARKGVGKIGRDVILVASIGGVVAVLMGAS

```


AUTHORS
Stover, C.K., Pham, X.Q., Erwin, A.L., Mizoguchi, S.D., Warren, P., Hickey, M.J., Brinkman, F.S., Hufnagle, W.O., Kowalik, D.J., Lagrou, M., Gardner, R.L., Goltzy, L., Tolentino, E., Westbrook-Wadman, S., Yuan, Y., Brody, L.L., Coulter, S.N., Folger, K.R., Kas, A., Lardig, K., Lim, R., Smith, K., Spencer, D., Wong, G.K., Wu, Z., and Paulsen, I.T.
Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen
Nature 406 (6799), 959-964 (2000)

JOURNAL
20437337

MEDLINE
2 (bases 1 to 12829)

REFERENCE
Stover, C.K., Pham, X.Q., Erwin, A.L., Mizoguchi, S.D., Warren, P., Hickey, M.J., Brinkman, F.S., Hufnagle, W.O., Kowalik, D.J., Lagrou, M., Gardner, R.L., Goltzy, L., Tolentino, E., Westbrook-Wadman, S., Yuan, Y., Brody, L.L., Coulter, S.N., Folger, K.R., Kas, A., Lardig, K., Lim, R., Smith, K., Spencer, D., Wong, G.K., Wu, Z., Paulsen, I.T., Reizer, J., Sater, M.H., Hancock, R.E.W., Lory, S. and Olson, M.V.
Direct Submission
Submitted (16-MAY-2000) Department of Medicine and Genetics, University of Washington Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA

TITLE
JOURNAL

FEATURES
source
1..12829
/organism="Pseudomonas aeruginosa"
/strain="PA01"
/db_xref="taxon:287"
219..947
/gene="PA0102"
219..947
/gene="PA0102"
/codon_start=1
/transl_table=11
/product="probable carbonic anhydrase"
/protein_id="AAG03492.1"
/db_xref="GI:9945929"
/translation="MPDRMGVKSQSDPEESADALKRIVDFQHFREVEPEQALF
KLLANSORPRAEITCADSRIVPELTQSSPDLEFTRNGVNPVPGMGNGVSTAI
EYAVIALGVHHTICGSLGDCGMAVAVLDPPLEPRMPYKALRHAEEVARTVAVNDCD
GASHPTLVTFRENVVADLDHFRPSPASASRLASQQLIHGMVYDIESAQIRAVDAKO
GRPLDGSHPVPAKPRYLS"
1158..2729
/gene="PA0103"
1158..2729
/gene="PA0103"
/codon_start=1
/transl_table=11
/product="probable sulfate transporter"
/protein_id="AAG03493.1"
/db_xref="GI:9945930"
/translation="MHSVFPSPSLRDLTPRDLMAVVVEVLVAPLCLMGIAIASGMPA
KGLIINGVIGVFLAGSPLOVSGPAGLAVELRYTGAVMLGPIILLAGAIOL
LAGRLRIGQWFRVTSFPAVYVGMGLAGITLIYVLSOHWMLDAPKASGIDLLAPQAA
FAALSLGMDSGDALGLGTAYMGMDKIRPRLRPLRALLGVSLLATLASMLA
LDVRREVANLGEAIWMLRPDLADLPDSLLAAVVAFTASAEILLSAVAVRHL
DPPSRMDRELISAGCVGNMLCGILGALPMTGIVYSSANVANGATRSALFPHGLML
AFVLLGSLRLIIPVASLAGVLYTVGVKLVDFKALGNLSRYGRMPLVYATLAIIV
TLLTGIVGHPALTLTKLVTKAARLKILARYGHEAEELRSGAATFLKVALSVLD
EYKPGTTLHVPMDNISYVDHACMELLEDMGRNAPQGSLSVIERALKRRLREGRLGS
VGLGARRNGAASPC"
2866..3462
/gene="PA0104"
2866..3462
/gene="PA0104"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAG03494.1"
/db_xref="GI:9945931"
/translation="MVDGFVQPVGEGLADQLVGHQVGYGVQVQDCLAENVDAVYVH
FHQAVRGVGVGVGVAAYVAGBORRGEELGDIQHLRDLPLDAGFAGHAGVAGVY
RVLFDYGVDEAGEYVRVIALRLDGMQVQAGAGQISQVTAAGFAEYVLLDEPFEES
GVHGVSHTTGMVKKRRNINAGARRLRDRDHNRFVYVNG"
3726..4850

gene
/gene="coxB"
/note="PA0105"
3726..4850
/gene="coxB"
/codon_start=1
/transl_table=11
/product="cytochrome c oxidase, subunit II"
/protein_id="AAG03495.1"
/db_xref="GI:9945932"
/translation="MLRHPVMMGFLLSAISQANAAMTVNMAPGATEVRSVFDLM
TIRICVIGLVFGAMFMSLVYHNRSTGQAPAHNRESTVEILMPYFVLVYMAV
PATRLIHLIVSEPELDVQVGYQWQYKYGQVETESLAPPODINRQAKDE
HYLEVEDPELVLPVSGTKRFLITSSDVHSMWVPAVAKRDAIPGVNDAWKVDEP
IYRQCAELCGKDHGEMPIYVYVVKPAAEEDQMLAKKEBAKVELTSEKMEBLVA
RGDKVYHTTCAACGACGAGMPMPALGSKIVTGPKEHLLEVNGVGTAMAFG
KQLEVDIAAVITYERNANGNDGQDVTFKDVAVAKQKQ"
4860..6452
/gene="coxA"
/note="PA0106"
4860..6452
/gene="coxA"
/codon_start=1
/transl_table=11
/product="cytochrome c oxidase, subunit I"
/protein_id="AAG03496.1"
/db_xref="GI:9945933"
/translation="MSAVIDTDPDHAGDHHNGPAGKLMRVLTNNHDKIGTLVLMFSF
MPLFGSSNMYVRAELFQPGQIIVPAFNNQMTWGLIIVGAVNPAFVGLANMI
PLMIGADNALPRNNNSFVLLPAAAGLLVSTLFMGCGGPNCGMPYAPLSTFAPHS
VTEFIFAIHLAGISSIMGATINVTATLNLRAPOMTLMKRLPLVWTLITAFLLIAMP
VLAVVYTMMDLHFGTSFSAAGGDPVLFQHVFWFCHPEVYIMLIPAFASVSLI
PTFAKPLFEGYTSNMYATIASIAFLSVVNAHMFVGVIGVTELFPMVATMLIAVPTG
KVFNWYTTMECSLTFEPEMLFAVAVFVILITGSGSGLMLIAPADFOYHDTYEVVA
HFVYLVPGAIFGIPASAYXVYLPKMGHMDVETLGLHFMMSIGNLAFFPMHFEVL
AGPRRIPDPYNLOPADPNMYSSIGAFMGITQLLFLIYIKIRGSKPAPAPMDGAE
GLEMSLSPAPVHTFSTPREV"
6463..7017
/gene="PA0107"
6463..7017
/gene="PA0107"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAG03497.1"
/db_xref="GI:9945934"
/translation="MSDAKYDTRRLVGRLLVTVLMFAFGFALVPLVDWCALGTING
KTAGSAYSGEOVDYGVREKVVQFMTSNIDMWEFSPSAGDQLVVHFGAVNOMVYARN
PSDKPRMRAIIPSTAPAEAAVYHKTDECFTQGVLPQGSIMPVRFYIDRDLPDV
RHVLAITLFDITARKPPVYVAGR"
7028..7915
/gene="coiII"
/note="PA0108"
7028..7915
/gene="coiII"
/codon_start=1
/transl_table=11
/product="cytochrome c oxidase, subunit III"
/protein_id="AAG03498.1"
/db_xref="GI:9945935"
/translation="MASHENYVPAOSKMPILIASIGLVTVFGLGTWENDLPAHGRHS
HGPIVTFVGGILITAYMLFGWGNVIRSRAGLISAOADRFSRGMSEFTSEVFEAA
FEGALFTVRRHFGWLGEGKAGVANHLMRNFOYSRPLLOTPOPKLIPPSAYIEPDK
LPLINTLILVTSSTVTFANHAKLNRRGPKAMALITVLGLTAFLILOAEETVYHYN
ELIGLTGAGIYGSFTFMITGFHGAHVTLGALILGIMLIRLRGHFDEHREHFEASW
YMHFVDVWIGLFEVYVI"
complement(7931..8140)
/gene="PA0109"
complement(7931..8140)
/gene="PA0109"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAG03499.1"


```

Query Match          34.7%; Score 30.2; DB 9; Length 714;
Best Local Similarity 60.2%; Pred. No. 1.2e+02;
Matches 50; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

OY 4 tgcagcttcctgcgtgcttgcctgcgacgcgcacccgcgcgcgagtcgagtcggcaaacg 63
    ||| ||||| ||||| ||||| ||||| |||
Db 569 TCCACATTCCTCTTGTGTGCTCCACCGCCACCACTACTACACCACTTCACATGCAGACCAT 628
    ||| ||||| ||||| ||||| ||||| |||

OY 64 caacgcgcttcagagatgccc 86
    ||| ||||| ||||| |||||
Db 629 CACTTCGGCTTCGAAGCCGCCG 651

RESULT 13
AC019055 176383 bp DNA PRI 03-JUL-2001
LOCUS Homo sapiens clone RP11-111P18, complete sequence.
AC019055
AC019055
AC019055.5 GI:14589677
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 176383)
Waterston,R.H.
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 176383)
Waterston,R.H.
Direct Submission
Submitted (30-DEC-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
3 (bases 1 to 176383)
Waterston,R.H.
Direct Submission
Submitted (09-AUG-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 176383)
Waterston,R.H.
Direct Submission
Submitted (10-SEP-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 176383)
Waterston,R.
Direct Submission
Submitted (30-SEP-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
6 (bases 1 to 176383)
Waterston,R.H.
Direct Submission
Submitted (03-JUL-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Jul 3, 2001 this sequence version replaced gi:9755483.

COMMENT
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site:http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H.NH0111P18
-----
Location/Qualifiers
1. 176383
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-111P18"

```


51003 68891 contig of 17889 bp in length
 68992 84339 contig of 15348 bp in length
 84440 97577 contig of 13138 bp in length
 97678 109261 contig of 11584 bp in length
 109362 118150 contig of 8789 bp in length
 118251 127677 contig of 9427 bp in length
 127778 134661 contig of 6884 bp in length
 134762 139706 contig of 4945 bp in length
 139807 145253 contig of 5447 bp in length
 145354 150450 contig of 5097 bp in length
 150551 156033 contig of 5483 bp in length
 156134 160013 contig of 3880 bp in length
 160114 160113: gap of 100 bp in length
 163444 163443: contig of 3330 bp in length
 163544 163543: gap of 100 bp in length
 166188 166287: gap of 100 bp in length
 166288 169598: contig of 3311 bp in length
 169599 169698: gap of 100 bp in length
 171442 171442: contig of 1744 bp in length
 171443 171542: gap of 100 bp in length
 171543 173750: contig of 2208 bp in length
 173751 173850: gap of 100 bp in length
 173851 175411: contig of 1561 bp in length

NOTE: This is a 'working draft' sequence. It currently consists of 28 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 29465: contig of 29465 bp in length
 29466 29565: gap of 100 bp
 29566 50902: contig of 21337 bp in length
 50903 51002: gap of 100 bp
 51003 68891: contig of 17889 bp in length
 68892 68991: gap of 100 bp
 68992 84339: contig of 15348 bp in length
 84340 84439: gap of 100 bp
 84440 97577: contig of 13138 bp in length
 97578 97677: gap of 100 bp
 97678 109261: contig of 11584 bp in length
 109262 109361: gap of 100 bp
 109362 118150: contig of 8789 bp in length
 118151 118250: gap of 100 bp
 118251 127677: contig of 9427 bp in length
 127678 127777: gap of 100 bp
 127778 134661: contig of 6884 bp in length
 134662 134761: gap of 100 bp
 134762 139706: contig of 4945 bp in length
 139707 139806: gap of 100 bp
 139807 145253: contig of 5447 bp in length
 145254 145353: gap of 100 bp
 145354 150450: contig of 5097 bp in length
 150451 150550: gap of 100 bp
 150551 156033: contig of 5483 bp in length
 156034 156133: gap of 100 bp
 156134 160013: contig of 3880 bp in length
 160014 160113: gap of 100 bp
 160114 163443: contig of 3330 bp in length
 163444 163543: gap of 100 bp
 163544 166187: contig of 2644 bp in length
 166188 169598: contig of 3311 bp in length
 169599 169698: gap of 100 bp
 171442 171442: contig of 1744 bp in length
 171443 171542: gap of 100 bp
 171543 173750: contig of 2208 bp in length
 173751 173850: gap of 100 bp
 173851 175411: contig of 1561 bp in length

FEATURES

SOURCE

175412 175511: gap of 100 bp
 175512 177277: contig of 1766 bp in length
 177278 177377: gap of 100 bp in length
 177378 178748: contig of 1371 bp in length
 178749 178848: gap of 100 bp
 178849 179788: contig of 1130 bp in length
 179789 180078: gap of 100 bp
 180079 181125: contig of 1047 bp in length
 181126 181225: gap of 100 bp
 181226 181791: contig of 566 bp in length
 181792 181881: gap of 100 bp
 181882 182903: contig of 1012 bp in length
 182904 183003: gap of 100 bp
 183004 184021: contig of 1018 bp in length
 184022 184121: gap of 100 bp
 184122 185147: contig of 1026 bp in length.

Location/Qualifiers
 1. 185147
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="18"
 /map="18q23"
 /clone="RP11-675M6"
 1. 29465
 /note="assembly-fragment"
 29566. 50902
 /note="assembly-fragment"
 51003. 68891
 /note="assembly-fragment clone_end:r7 vector_side:right"
 68992. 84339
 /note="assembly-fragment"
 84440. 97577
 /note="assembly-fragment"
 97678. 109261
 /note="assembly-fragment"
 109362. 118150
 /note="assembly-fragment"
 118251. 127677
 /note="assembly-fragment"
 127778. 134661
 /note="assembly-fragment"
 134762. 139706
 /note="assembly-fragment"
 139807. 145253
 /note="assembly-fragment"
 145354. 150450
 /note="assembly-fragment"
 150551. 156033
 /note="assembly-fragment"
 156134. 160013
 /note="assembly-fragment"
 160114. 163443
 /note="assembly-fragment"
 163544. 166187
 /note="assembly-fragment"
 166288. 169598
 /note="assembly-fragment"
 169699. 171442
 /note="assembly-fragment"
 171543. 173750
 /note="assembly-fragment"
 173851. 175411
 /note="assembly-fragment"
 175512. 177277
 /note="assembly-fragment"
 177378. 178748
 /note="assembly-fragment"
 178849. 179788
 /note="assembly-fragment"
 180079. 181125
 /note="assembly-fragment"
 181226. 181791
 /note="assembly-fragment clone_end:SP6 vector_side:right"

Thu Jan 10 08:15:52 2002

us-09-712-768-7.rge

Page 18

```
misc_feature 181892..182903
              /note="assembly-fragment"
```

Query Match	34.3%	Score 29.8;	DB 2;	Length 185147;
Best Local Similarity	63.0%;	Pred. No. 63;		
Matches 46;	Conservative	0;	Mismatches 27;	Indels 0;
				Gaps 0;

Oy 1 gattcgatcttcctgctgcgtcctgatccgcacacctcgcggtagtgatgtcgycnaaa 60
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 70885 GATCCATCATCAGAGTGTCTGCTTCCTTCGGGCTCCAGGGCAATCCCACATGCTGC AAAA 70944

```

Qy      61 cagcaccgtcggt 73
      || | | | |
Db 70945 CATAGCTTCTGCT 70957

```

Search completed: January 10, 2002, 00:57:09
Job time: 3117 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 10, 2002, 00:59:21 ; Search time 129.27 Seconds

(without alignments)
576.988 Million cell updates/sec

Title: US-09-712-768-7

Perfect score: 1 gattcgattcttcctgtcgtgt.....tcggttcgagatgcgcga 87

Sequence:

IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

```

1: N_Geneseq_1101.*
2: /SIDS2/gcgdata/geneseq/geneseqn/NA1980.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseqn/NA1981.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseqn/NA1982.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseqn/NA1983.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseqn/NA1984.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseqn/NA1985.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseqn/NA1986.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseqn/NA1987.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseqn/NA1988.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseqn/NA1989.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseqn/NA1990.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseqn/NA1991.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseqn/NA1992.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	87	100.0	87	AAA91492	G. oxydans cytochr
2	29.6	34.0	1837	AAH22961	Human phosphatase
3	27.4	31.5	23666	AAO10190	Cephalosporin anti
4	27	31.0	956	AAO105907	Mitochondrial cyto
5	26.8	30.8	373	AAA67153	Eucalyptus grandis
6	26	29.9	613	AAH35339	Human colon cancer
7	26	29.9	1227	AAH52068	Mycobacterium tube
8	26	29.9	2505	AAH94687	naph gene encoding
9	26	29.9	2765	AAH41550	Human Rtl gamma n
10	26	29.9	3131	AAH41549	Human Rtl beta nu
11	26	29.9	3332	AAH41548	Human Rtl alpha'

C	12	26	29.9	3344	22	AAH41505	Human Rtl1 encodin
C	13	25.8	29.7	1860	21	AAH265292	Human secreted pro
C	14	25.6	29.4	599	20	AAH228115	Human astrocyte ab
C	15	25.6	29.4	611	20	AAH228114	Human astrocyte ab
C	16	25.6	29.4	690	20	AAH84123	Human CD27 binding
C	17	25.6	29.4	885	20	AAH84122	Human CD27 binding
C	18	25.6	29.4	12152	22	AAH08699	M. catibonacea DNA
C	19	25.4	29.2	116	22	AAH04692	Gene expression pr
C	20	25.4	29.2	183	22	AAH04705	Gene expression pr
C	21	25.4	29.2	260	20	AAH87886	EST clone FC274.
C	22	25.4	29.2	577	21	AAH69796	Human breast tumo
C	23	25.4	29.2	637	22	AAH64221	Human secreted pro
C	24	25.4	29.2	728	22	AAH62206	Human prostate can
C	25	25.4	29.2	818	19	AAH61271	3' cDNA sequence o
C	26	25.4	29.2	818	19	AAH58483	Human immunogeni
C	27	25.4	29.2	818	21	AAH06246	Human prostate tum
C	28	25.4	29.2	818	22	AAH10005	Human prostate spe
C	29	25.4	29.2	818	22	AAH93362	Human prostate spe
C	30	25.4	29.2	818	22	AAH84676	Human prostate spe
C	31	25.4	29.2	818	22	AAH02427	Human prostate spe
C	32	25.4	29.2	873	20	AAH28117	Human prostate spe
C	33	25.4	29.2	923	21	AAH63445	Human astrocyte ab
C	34	25.4	29.2	926	19	AAH53092	Human secreted pro
C	35	25.4	29.2	954	22	AAH68881	Human cytochrome c
C	36	25.4	29.2	990	21	AAH59763	Mitochondrial cyto
C	37	25.4	29.2	16569	20	AAH21492	Human secreted pro
C	38	25.4	29.2	16569	20	AAH21493	Mutated human mito
C	39	25.4	29.2	16569	20	AAH21491	Mutated human mito
C	40	25.4	29.2	16569	21	AAH57203	Human mitochondria
C	41	25.2	29.0	564	21	AAH39592	M. tuberculosis an
C	42	25.2	29.0	1183	20	AAH24104	Human endometrium
C	43	25.2	29.0	1754	22	AAH81366	Quorum sensing con
C	44	25.2	29.0	2153	22	AAH93904	Human cDNA encodin
C	45	25.2	29.0	2197	21	AAH98120	Human colon cancer

ALIGNMENTS

RESULT	1	
AAA91492		
ID	AAA91492 standard; DNA; 87 BP.	
AC	AAA91492;	
XX		
DT	06-AUG-2001 (first entry)	
DE		
XX		
XX	G. oxydans cytochrome C oxidase (COII) coding sequence fragment.	
KW	Cytochrome C oxidase; COI; COII; COIII; COI complex; D-sorbitol;	
KW	oxidative fermentation; electron transfer; respiratory chain; L-sorbose;	
KW	2-Keto-L-gluconic acid; 2KGA production; aldehyde production;	
KW	carboxylic acid production; ketone production; ds.	
OS		
XX		
XX	Gluconobacter oxydans.	
FT		
FT	Key	Location/Qualifiers
FT	CDS	1..87
FT		/*tag= a
FT		/product= "COII"
PN	EP1103603-A2.	
XX		
PD	30-MAY-2001.	
XX		
PF	14-NOV-2000; 2000EP-0124785.	
XX		
PR	17-NOV-1999; 99EP-0122842.	
XX		
PA	(HOFF) HOFFMANN LA ROCHE & CO AG F.	
XX		
PI	Asakura A, Hoshino T, Shinjoh M;	
XX		

DR MPI: 2001-357953/38.
DR P-PSDB: AAY97753.
XX
XX New cytochrome c oxidase complex having cytochrome c oxidase activity
PT from Gluconobacter oxydans DSM 4023, useful in mediating electron
PT transfer in respiratory chain or producing 2-keto-L-gulononic acid from
PS L-sorbose or D-sorbitol
PS
PS Claim 15: Page 26; 42pp; English.
XX
XX This sequence encodes a fragment of the Gluconobacter oxydans
CC cytochrome c oxidase, COI11. The invention relates to a cytochrome c
CC cytochrome COI complex. The COI complex is useful in improving oxidative
CC fermentation and is an essential component mediating electron transfer
CC in the respiratory chain. The recombinant microorganism and the
CC cytochrome c oxidase may be used in the genetic preparation of a
CC recombinant COI complex and in the production of 2-Keto-L-gluconic acid
CC (2KGA) from L-sorbose or D-sorbitol in a culture medium. The COI is also
CC useful as a terminal oxidase, oxidising cytochrome c, an electron
CC acceptor from an enzyme belonging to dehydrogenase for the production of
CC aldehydes, carboxylic acids and ketones from alcohols and aldehydes,
CC especially the production of 2KGA from L-sorbose or D-sorbitol.
XX
XX Sequence 87 BP; 14 A; 26 C; 26 G; 21 T; 0 other;
SO

Query Match	100.0%	Score 87	DB 22	Length 87
Best Local Similarity	100.0%	Pred. No. 3,76-20		
Matches	87	Conservative	0	Mismatches 0; Indels 0; Gaps 0
QY	1	gatttgatcttctctctgtctgtctgataccgacatctctgcggtgcggtatgcgcaaaa	60	
Db	1	gatttgatcttctctctgtctgtctgataccgacatctctgcggtgcggtatgcgcaaaa	60	
QY	61	cagcacgttcgtcttcgagaatgcgcgca	87	
Db	61	cagcacgttcgtcttcgagaatgcgcgca	87	

P1 Bandman O., Tang YF, Azimzal Y, Yue H, Baughn MR, Hillman JL;
PI Lal P, Wang E, Gandhi AR, Policky JU, Macher P;
XX WPI: 2001-451905/48.
DR P-PSDB; AAB85358.
XX
PT Novel phosphatases useful for treating diseases associated with
PT decreased expression of functional phosphatases, e.g., Alzheimer's
PT disease, Huntington's disease, immune disorders, and cancers -
XX
XX
XX Claim 5; Page 100; 103pp; English.
XX
XX The invention provides human phosphatases (PP) and polynucleotides
CC encoding the phosphatases. The polypeptides can be expressed by standard
CC recombinant methodology. The PP are useful for treating a disease or
CC condition associated with decreased expression of functional
CC phosphatases. Compositions containing agonists or antagonists of PP may
CC be used to treat a disease associated with decreased expression or
CC overexpression of PP, respectively. Such diseases may include
CC Alzheimer's disease, Huntington's disease, immune disorders, and
CC cancers. The present sequence represents a human phosphatase encoding
CC cDNA.
XX
XX Sequence 1837 BP; 392 A; 527 C; 555 G; 363 T; 0 other;

	Query Match	34.0%;	Score 29.6;	DB 22;	Length 1837;
	Best Local Similarity	61.8%;	Pred. No. 1;		
Matches	47; Conservative	0;	Mismatches	29;	Indels 0; Gaps
Oy	7 atcttcgtgcgagcgtccatcgcatctctggcgtgatgacgaagaaggaaac	66			
Dd	1128 ATCTCCTGTGACGCACAGTCGGTCCATCTCACACACTGTTGGCAACAAGCCA	1065			
Oy	67 gtcsgtttcgagatatg 82				
Dd	1068 GAAGGGCTTGAGCGTGG 1053				

RESULT	2	
AAH22961/c		
ID	AAH22961	standard; cDNA, 1837 BP.
XX		
AC	AAH22961;	
XX		
DT	17-SEP-2001	(first entry)
XX		
DE	Human phosphatase (PP) encoding cDNA (clone ID 1723447CB1).	
XX		
XX	Phosphatase; PP; Alzheimer's disease; Huntington's disease; human;	
KW	immune disorder; cancer; neotrophic; immunomodulatory; cyostatic;	
XX	neuroprotective; anticonvulsant; ss.	
KX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	1..1323
FT		/*tag= a
FT		/product= "phosphatase"
XX		
FN	W0200153469-A2.	
XX		
PD	26-JUL-2001.	
XX		
PF	18-JAN-2001; 2001WO-US02088.	
XX		
PR	21-JAN-2000; 2000US-0177719.	
PR	28-JAN-2000; 2000US-0178988.	
PR	25-FEB-2000; 2000US-0184959.	
PR	17-MAR-2000; 2000US-0190142.	
XX		
PA	(INCY-)	INCYTE GENOMICS INC.
XX		

RESULT	3
AAQ10190	
ID	AAQ10190 standard; DNA, 23666 BP.
XX	
AC	AAQ10190;
XX	
DT	27-MAR-1991 (first entry)
XX	
DE	Cephalosporin antibiotic biosynthetic genes.
XX	
KW	cephalosporin; antibiotic;
KW	S-(L-alpha-aminoadipyl)-L-cysteinyl-D-, valine synthetase;
KW	isopenicillin N synthetase; isopenicillin N epimerase;
KW	deacetoxycephalosporin C synthetase; beta-lactamase;
KW	deacetoxycephalosporin C hydroxylase; ss.
XX	
OS	Lysobacter lactamgenus.
XX	
PH	Location/Qualifiers
FT	5524..16692
FT	/*tag= a
FT	/label= ORF 1
FT	16761..17741
FT	/*tag= b
FT	/label= ORF 2
FT	17802..18761
FT	/*tag= c
FT	/label= ORF 3
FT	18798..19739
FT	/*tag= d
FT	/label= ORF 4
FT	19802..21061
FT	/*tag= e
FT	/label= ORF 5

XX	WPI: 2000-686079/67.
DR	P-PSDB: AAB26917.
XX	
PT	New nucleic acid encoding bacterial nitric oxide reductase, useful e.g.
PT	for preparing transgenic organisms able to detoxify wastes
XX	
PS	Example 2; Columns 31-34; 24pp; English.
XX	
CC	The present sequence is the nupA gene from <i>Pseudomonas</i> sp. strain G-179.
CC	This sequence encodes the large subunit of periplasmic nitrate reductase.
CC	This sequence can be used to create transgenic organisms that can
CC	denitrify toxic waste. In addition, the present sequence may be used to
CC	produce probes and primers for identification of new denitrifying
CC	bacteria.
XX	
SO	Sequence 2505 BP; 557 A; 789 C; 712 G; 447 T; 0 other;
QY	30 ccgactcctgcgcggtgcgagtgcgcaaacagcagcgtcggttcgagatggccgca 87
DB	1575 cctgacctgcgcgcgagatgtggtgtaaaagaagcgtctacgcaatgccgaa 1632
XX	
RESULT	9
AAH41550/c	
ID	AAH41550 standard; DNA; 2765 BP.
XX	
AC	AAH41550;
XX	
DT	14-SEP-2001 (first entry)
XX	
DE	Human Rlt1 gamma nucleotide sequence.
XX	
KW	Mouse; human; combined DNA/RNA molecule; Rlt1; tumour suppressor;
KW	2-3 type zinc finger structure; cancer; diagnosis; carcinogenesis;
KW	gene therapy; ds.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	Location/Qualifiers
FT	CD5
FT	11..2116
FT	/*tag= a
FT	/product= "Rlt1 gamma"
XX	
PN	MO200132859-A1.
XX	
PD	10-MAY-2001.
XX	
PF	14-JUL-2000; 2000WO-JP04765.
XX	
PR	29-OCT-1999; 99JP-0310420.
XX	
PA	(MOCH) MOCHIDA PHARM CO LTD.
XX	
PI	Kominami R;
XX	
DR	WPI: 2001-316438/33.
DR	P-PSDB: AAB99341.
XX	
PT	New zinc finger protein and gene encoding it for detecting and
PT	diagnosing cancer, estimating the risk of carcinogenesis, and for gene
XX	therapy
XX	
PS	Claim 40; Page 81-84; 119pp; Japanese.
XX	
CC	The present invention describes a combined DNA/RNA molecule designated
CC	Rlt1, which has a 2-3 type zinc finger structure and tumour suppressor
CC	activity. Rlt1 has cytosstatic activity and can be used in gene therapy.

XX PN W09947692-A2.
XX PD 23-SEP-1999.
XX PF 18-MAR-1999; 99WO-US05930.
XX PR 18-MAR-1998; 98US-0040736.
XX PA (UYO) UNIV JOHNS HOPKINS.
XX PI Rothstein JD, Lin CG, Bristol LA;
XX DR WPI: 1999-580308/49.
XX PT Use of glutamate transporter 2 mRNA for developing products for
XX PT diagnosis or treatment of neurological disorders, e.g. amyotrophic
XX PT lateral sclerosis, Huntington's, Parkinson's or Alzheimer's disease -
XX PS Claim 11; Fig 4; 99pp; English.
XX CC This sequence represents aberrant glutamate transporter 2
CC (excitatory amino acid transporter 2, EAAT 2) cDNA #4, derived from
CC human astrocyte mRNA. This sequence comprises part of the 5'
CC untranslated region of the EAAT 2 gene and a portion of exon 8,
CC plus exons 9-11, of the wild-type EAAT 2 cDNA (AA282111). mRNA
CC encoding aberrant forms of EAAT 2 are abnormally spliced variants of the
CC wild-type sequence and contain an intron sequence (AA282112) or are
CC missing one or more exons (AA28113-228121). At least one type of
CC aberrant EAAT 2 mRNA was detected in a majority of patients having, or
CC suspected of having, certain neurological disorders associated with
CC excitotoxicity which affect motor neuron function. These disorders
CC include amyotrophic lateral sclerosis (ALS) Huntington's disease,
CC Parkinson's disease and Alzheimer's disease. The aberrant mRNA, or cDNA
CC derived from it, can be used as a marker for these diseases in a novel
CC method of detection of a neurological disorder. This method comprises
CC obtaining a biological sample from a patient, and characterising the EAAT
CC 2 mRNA. The presence of at least one aberrant EAAT 2 mRNA in the sample
CC is indicative of the neurological disorder in the patient. The methods
CC can be used for detecting neurological disorders such as ALS,
CC Huntington's disease, Parkinson's disease or Alzheimer's disease. The
CC methods can also be used for monitoring the efficiency of therapy and for
CC identifying compounds for the treatment of neurological disorders.
CC Recombinant vectors comprising antisense nucleic acids homologous to the
CC aberrant sequences can be used for treating or preventing neurological
CC disorders. The methods provide sensitive and reliable methods for
CC detecting specific neurological disorders in living patients with minimal
CC impact to the nervous system. In addition, the methods can serve as an
CC indicator of predisposition for or susceptibility to the disorders.

XX S0 Sequence 611 BP; 153 A; 174 C; 162 G; 122 T; 0 other;

Query Match: 29.4%; Score 25.6; DB 20; Length 611;
Best Local Similarity 59.7%; Pred. No. 18;
Matches 43; Conservative 0; Mismatches 29; Indels 0; Gaps 0

OY 15 gctggtcgtgcctgatccgcacatcctcgcgcggtgcatgttcggcaaacagcacgttctt 74
DB 242 gctgtgccacatgctctcatcttgacagcgcgtggcgccctgccaagaagaaatcagctt 301
OY 75 cgagatggcgcg 86
DB 302 gctgtgtgctgc 313

Search completed: January 10, 2002, 00:59:23
Job time: 3176 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd

OM nucleic - nucleic search, using sw model

Run on: January 10, 2002, 01:00:12 ; Search time 49.55 Seconds

397.651 million cell updates/sec

Title: US-09-712-768-7

Sequence: 1 gattcgatcttctctgtgt.....tcggtttcgaagatgcccga 87

Scoring table: IDENTITY_NUC

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

```

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

```

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued_Patents_NA:*

```

1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	26.6	30.6	4403765	4	US-09-103-840A-2	Sequence 2, Appli
2	26.6	30.6	4411529	4	US-09-103-840A-1	Sequence 1, Appli
3	26	29.9	2505	3	US-09-354-129-9	Sequence 9, Appli
4	25.6	29.4	690	3	US-08-865-297-3	Sequence 3, Appli
5	25.6	29.4	885	3	US-08-865-297-1	Sequence 1, Appli
C	25.4	29.2	818	4	US-09-020-935-6	Sequence 6, Appli
C	25.4	29.2	818	4	US-09-030-607-6	Sequence 6, Appli
C	25.4	29.2	818	4	US-08-219-842-3	Sequence 3, Appli
7	25.4	29.2	856	1	US-08-451-096-3	Sequence 3, Appli
8	25.4	29.2	856	1	US-08-810-599-82	Sequence 82, Appli
9	25.4	29.2	926	2	US-08-413-740A-3	Sequence 3, Appli
10	25.4	29.2	954	4	PCT-US95-04053-3	Sequence 3, Appli
11	25.4	29.2	954	5	US-09-097-889-2	Sequence 2, Appli
12	25.4	29.2	16569	4	US-08-140-729A-6	Sequence 6, Appli
13	25.4	28.3	1800	1	US-08-546-666-6	Sequence 6, Appli
14	24.6	28.3	1800	2	US-08-916-745-6	Sequence 6, Appli
15	24.6	28.3	1800	2	US-08-948-569A-3	Sequence 3, Appli
16	24.6	28.3	1800	2	US-08-663-808-5	Sequence 5, Appli
17	24.6	28.3	1800	2	US-09-042-928-6	Sequence 6, Appli
18	24.6	28.3	1800	2	US-08-546-661-6	Sequence 6, Appli
19	24.6	28.3	1800	2	US-09-042-960-6	Sequence 6, Appli
20	24.6	28.3	1800	2	US-09-198-650-6	Sequence 6, Appli
21	24.6	28.3	1800	3	US-09-332-740-5	Sequence 5, Appli
22	24.6	28.3	1800	3	US-09-042-913-6	Sequence 6, Appli
23	24.6	28.3	1800	3	US-09-188-449-5	Sequence 5, Appli
24	24.6	28.3	1800	3	US-09-042-937-6	Sequence 6, Appli
25	24.6	28.3	1800	3	US-09-042-937-6	Sequence 6, Appli
26	24.6	28.3	1800	3	US-09-042-937-6	Sequence 6, Appli
27	24.6	28.3	1800	3	US-09-042-937-6	Sequence 6, Appli

28	24.6	28.3	1800	4	US-09-397-232A-3	Sequence 3, April
29	24.2	27.8	264	2	US-08-774-028A-2	Sequence 2, April
30	24.2	27.8	264	4	US-09-244-093-2	Sequence 2, April
31	24.2	27.8	1197	3	US-08-890-713-7	Sequence 7, April
32	24.2	27.8	1230	3	US-08-890-713-1	Sequence 7, April
33	24.2	27.8	1230	3	US-08-890-713-1	Sequence 8, April
34	24.2	27.8	1262	3	US-08-890-713-37	Sequence 3, April
35	24.2	27.8	2019	4	US-09-063-950-3	Sequence 3, April
36	24.2	27.8	2852	4	US-09-063-950-1	Sequence 1, April
37	24	27.6	1860	3	US-09-071-434-1	Sequence 2, April
38	24	27.6	4411529	4	US-09-103-840A-2	Sequence 1, April
39	23.8	27.4	954	2	US-08-073-275-3	Sequence 3, April
40	23.8	27.4	1626	4	US-09-158-767-14	Sequence 14, April
41	23.8	27.4	2181	4	US-09-158-767-10	Sequence 10, April
42	23.6	27.1	135	5	PCT-US94-14467-1	Sequence 1, April
43	23.6	27.1	1694	1	US-08-698-551-9	Sequence 9, April
44	23.6	27.1	1694	2	US-08-602-228-9	Sequence 9, April
45	23.6	27.1	1694	2	US-08-494-040B-9	Sequence 9, April

ALIGNMENTS

RESULT 1

; Sequence 2, Application US/09103840A

GENERAL INFORMATION:

APPLICANT: WHITE, Owen R.

APPLICANT: VENTER, John C.

TITLE OF INVENTION: TUBERCULOSIS

CURRENT APPLICATION NUMBER: US/09/103,840A

NUMBER OF SEQ ID NOS: 2

[illegible]

TYPE: DNA

; FEATURE:

OTHER INFORMATION: "n" bases at various p

US-09-103-840A-2

0
1
2
3
4
5
6
7
8
9
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z

Best Local Sim

7

—
—
—
—
—

0
n
→

A gel electrophoresis image showing a single band in the lane labeled 'C'.

RESULT 2

; Sequence 1, Ap

; GENERAL INFORM

APPLICANT: W

APPLICANT: V

APPLICATION NUMBER: US/08/865,297
FILING DATE: 29-MAY-1997
CLASSIFICATION: 435


```

PRIOR APPLICATION DATA:
  APPLICATION NUMBER:
  FILING DATE:
  ATTORNEY/AGENT INFORMATION:
    NAME: Silveri, Jean M.
  REGISTRATION NUMBER: 39,030
  REFERENCE/DOCKET NUMBER: DEN-024
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (617) 227-7400
    TELEFAX: (617) 227-5941
  INFORMATION FOR SEQ ID NO: 1:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 885 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: single
      TOPOLOGY: linear
    MOLECULE TYPE: CDNA
    FRAGMENT TYPE: Internal
  FEATURE:
    NAME/KEY: CDS
    LOCATION: 133..699
    US-08-865-297-1
  
```

```

Query Match          29.4%; Score 25.6; DB 3; Length 885;
Best Local Similarity 57.5%; Pred. No. 4.2;
Matches 46; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
  
```

```

OY 7 atctctcgtgctgctcgtatccgcatccctgcggtgagatgctggaacagcac 66
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 248 AACCTCCCTGGCTGACGCGGACCTTGAGCTGTAGCGGGCCACGTCGGAAGGGCAG 189
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 67 gtccgttcgagatgagccgc 86
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 188 CTCGCTTGCGCATGGCCGC 169
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  
```

```

RESULT 6
US-09-020-956-6/c
; Sequence 6, Application US/09020956
; Patent No. 6261562
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillin, Davin C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS FO
; NUMBER OF SEQUENCES: 178
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/020,956
; FILING DATE: 09-FEB-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.427C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 818 base pairs
; TYPE: nucleic acid
  
```

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; US-09-020-956-6
  
```

```

Query Match          29.2%; Score 25.4; DB 4; Length 818;
Best Local Similarity 56.6%; Pred. No. 4.8;
Matches 47; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
  
```

```

OY 4 tcatctctcgtgctgctcgtatccgcatccctgcggtgagatgctggaacag 63
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 171 TCAACTTCTCCTCACTATGCTTCATCCGCGCACTAATATTTCACTTTATCAATCAAAACT 112
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 64 cactggttcgagatgagccgc 86
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 111 CACTTGGCTTCGAAGCCGCCGC 89
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  
```

```

RESULT 7
US-09-030-607-6/c
; Sequence 6, Application US/09030607
; Patent No. 6262245
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillin, Davin C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS
; NUMBER OF SEQUENCES: 224
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/030,607
; FILING DATE: 25-FEB-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.427C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 818 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; US-09-030-607-6
  
```

```

Query Match          29.2%; Score 25.4; DB 4; Length 818;
Best Local Similarity 56.6%; Pred. No. 4.8;
Matches 47; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
  
```

```

OY 4 tcatctctcgtgctgctcgtatccgcatccctgcggtgagatgctggaacag 63
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 171 TCAACTTCTCCTCACTATGCTTCATCCGCGCACTAATATTTCACTTTATCAATCAAAACT 112
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 64 cactggttcgagatgagccgc 86
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 111 CACTTGGCTTCGAAGCCGCCGC 89
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
  
```

```

1      RESULT      8
2      US-08-219-842-3
3      : Sequence 3, Application US/08219842
4      : Patent No. 5565323
5      : GENERAL INFORMATION:
6      : APPLICANT: Parker, W. D.
7      : APPLICANT: Herrnsdorf, Corinna
8      : TITLE OF INVENTION: Diagnostic and Therapeutic Compositions
9      : TITLE OF INVENTION: for Alzheimer's Disease
10     : NUMBER OF SEQUENCES: 95
11     : CORRESPONDENCE ADDRESS:
12     : ADDRESSEE: Campbell and Flores
13     : STREET: 4370 La Jolla Village Drive, Suite 700
14     : CITY: San Diego
15     : STATE: California
16     : COUNTRY: USA
17     : ZIP: 92122
18     :
19     : COMPUTER READABLE FORM:
20     : MEDIUM TYPE: Floppy disk
21     : COMPUTER: IBM PC compatible
22     : OPERATING SYSTEM: PC-DOS/MS-DOS
23     : SOFTWARE: Patentin Release #1.0, Version #1.25
24     : CURRENT APPLICATION DATA:
25     : APPLICATION NUMBER: US/08/219, 842
26     : FILING DATE: 30-MAR-1994
27     : CLASSIFICATION: 435
28     : ATTORNEY/AGENT INFORMATION:
29     : NAME: Campbell, Cathryn A.
30     : REGISTRATION NUMBER: 31,815
31     : REFERENCE/DOCKET NUMBER: P-AG 9504
32     : TELECOMMUNICATION INFORMATION:
33     : TELEPHONE: (619) 535-9001
34     : TELEFAX: (619) 535-8949
35     : INFORMATION FOR SEQ. ID NO.: 3:
36     : SEQUENCE CHARACTERISTICS:
37     : LENGTH: 856 base pairs
38     : TYPE: nucleic acid
39     : STRANDEDNESS: double
40     : TOPOLOGY: linear
41     :
42     : US-08-219-842-3

```

```

Query Match Similarity 29.2%: Score 25.4; DB 1; Length 856;
Best Local Similarity 56.6%: Pred. No. 4.9;
Matches 47; Conservative 0; Mismatches 36; Indels 0; Gaps 0.

QY      4  tgcagcttcctgctggtcgcctcagccagccatccgcgcgcggtgcgaltgcgcaaacag 63
      || | ||||| ||||| ||||| ||||| || | |||||
Db      704  TCAACTTTCCTCACTATCTGCTTTCATCGCCCAACTAATATTCTCACTTTACATCCAAACAT 763

QY      64  cacgctcgattcgcagatgagcgcg 86
      || | | || ||||| |||||
Db      764  CACTTGGCGCTTCGACAGCCGCCGC 786

RESULT      9
US-08-451-096-3
; Sequence 3, Application US/08451096
; Patent No. 5760205
GENERAL INFORMATION:
APPLICANT: Parker, W. D.
APPLICANT: HerrinStadt, Corinna
TITLE OF INVENTION: Diagnostic and Therapeutic Compositions
TITLE OF INVENTION: for Alzheimer's Disease
NUMBER OF SEQUENCES: 95
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122

```

```

1  COMPUTER READABLE FORM:
2  MEDIUM TYPE: Floppy disk
3  COMPUTER: IBM PC compatible
4  OPERATING SYSTEM: PC-DOS/MS-DOS
5  SOFTWARE: PatentIn Release #1.0, Version #1.25
6  CURRENT APPLICATION NUMBER: US/08/451,096
7  APPLICATION NUMBER: US
8  FILING DATE:
9  CLASSIFICATION: 435
10 PRIOR APPLICATION DATA:
11 APPLICATION NUMBER: US 08/219,842
12 FILING DATE: 30-MAR-1994
13 ATTORNEY/AGENT INFORMATION:
14 NAME: Campbell, Cathryn A.
15 REGISTRATION NUMBER: 31,815
16 REFERENCE/DOCKET NUMBER: P-AG 9504
17 TELECOMMUNICATION INFORMATION:
18 TELEPHONE: (619) 535-9001
19 TELEFAX: (619) 535-8949
20 INFORMATION FOR SEQ ID NO: 3:
21 SEQUENCE CHARACTERISTICS:
22 LENGTH: 856 base pairs
23 TYPE: nucleic acid
24 STRANDEDNESS: double
25 TOPOLOGY: linear
26 US-08-451-096-3

```

Query Match	29.28;	Score 25.4;	DB 1;	Length 856;
Best Local Similarly	56.68;	Pred. No. 4.9;		
Matches 47;	Conservative 0;	Mismatches 36;	Indels 0;	Gaps 0;

[illegible]

RESULT 10
 US-08-810-599-82
 Sequence 82. Application US/08810599
 Patent No. 5976798
 GENERAL INFORMATION:
 APPLICANT: PARKER, W. Davis
 APPLICANT: HERRNSTADT, Corinna
 APPLICANT: GHOSH, Soumitra S.
 APPLICANT: FAHY, Eoin
 TITLE OF INVENTION: Methods for Detecting Mitochondrial Mutations
 TITLE OF INVENTION: Diagnostic for Alzheimer's Disease and Methods for Determin
 TITLE OF INVENTION: of Mitochondrial Nucleic Acid
 NUMBER OF SEQUENCES: 82
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Kenyon & Kenyon
 STREET: 1025 Connecticut Avenue, N.W., Suite 600
 CITY: Washington
 STATE: D.C.
 COUNTRY: US
 ZIP: 20036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.25" Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Wordperfect 6.1 for Windows
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/810,599
 FILING DATE: Concurrent Herewith
 CLASSIFICATION: 436
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/757,438
 FILING DATE: 27 No. 5976798 1996

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/413,740P

```

```

; REGISTRATION NUMBER: 34297
; REFERENCE/DOCKET NUMBER: 2105/7

```

TELECOMMUNICATION INFORMATION:


```

; Sequence 2, Application US/08865297
; Patent No. 6010853
; GENERAL INFORMATION:
; APPLICANT: Prasad V.S. Kanteti, Zhaoxui Ao, and Stuart F.
; TITLE OF INVENTION: The Siva Genes, No. 6010853el Genes Involved in
; TITLE OF INVENTION: CD27-Mediated Apoptosis
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/865,297
; FILING DATE: 29-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Silveri, Jean M.
; REGISTRATION NUMBER: 39,030
; REFERENCE/DOCKET NUMBER: DFN-024
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 189 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-865-297-2

alignment_scores:
  Quality: 66.00      Length: 27
  Ratio: 3.000        Gaps: 0
  Percent Similarity: 81.481  Percent Identity: 40.741

alignment_block:
US-09-712-768-7/rev x US-08-865-297-2  ..

Align seg 1/1 to: US-08-865-297-2 from: 1 to: 189

86 GCGGCATCTCGAAGACGACGTGTTTGGCGACATCGCAGCGCGAG 37
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
13 AlaAlaMetProLysArgSerCysProPheAlaSpValAlaProLeuGl 29
:|||||:|||||:|||||:|||||:|||||:|||||:
36 GATCGGATCGACGACGACGACGACGACGACGACGACGACGACGAC 6
:|||||:|||||:|||||:|||||:|||||:|||||:
29 nleuLysValArgValSerGlnArgGluLeu 39

seq_name: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:US-09-097-889-19

seq_documentation_block:
; Sequence 19, Application US/09097889
; Patent No. 6218117
; GENERAL INFORMATION:
; APPLICANT: Herrstadt, Corrina
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Davis, Robert E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING
; TITLE OF INVENTION: AGENTS THAT QUANTITATIVELY ALTER DETECTABLE
; TITLE OF INVENTION: EXTRAMITOCHONDRIAL DNA: MITOCHONDRIAL DNA RATIOS

```

```

; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/097,889
; FILING DATE: 15-JUN-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rosenman Ph.D., Stephen J.
; REGISTRATION NUMBER: 43,058
; REFERENCE/DOCKET NUMBER: 660088.417
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 261 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-09-097-889-19

alignment_scores:
  Quality: 66.00      Length: 28
  Ratio: 3.300        Gaps: 0
  Percent Similarity: 71.429  Percent Identity: 50.000

alignment_block:
US-09-712-768-7 x US-09-097-889-19  ..

Align seg 1/1 to: US-09-097-889-19 from: 1 to: 261

4 TCGATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCGGCGTGCAGATGC 53
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
212 SerThrPheLeuThrIleCysPheIleArgGlnLeuMetPheHisPheTh 228
:|||||:|||||:|||||:|||||:|||||:|||||:
54 GCGAAGACGACGACGCTGCTTTCGAGATGCCGCGCA 87
:|||||:|||||:|||||:|||||:|||||:|||||:
228 rSerLysHisHisPheGlyPheGlnAlaAlaAla 239

seq_name: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:US-08-924-345-2

seq_documentation_block:
; Sequence 2, Application US/08924345
; Patent No. 6224878
; GENERAL INFORMATION:
; APPLICANT: LEUNG-TACK Patricia
; APPLICANT: LEGASTRELOIS Isabelle, Christine, Marie-Andree
; APPLICANT: AUDONNET Jean-Christophe, Francis
; APPLICANT: RIVIERE Michel, Emile, Albert
; TITLE OF INVENTION: Mutants and vaccines of the Infectious
; TITLE OF INVENTION: Bovine Rhinotracheitis virus
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LARSON AND TAYLOR
; STREET: 727 SOUTH TWENTY-THIRD STREET
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```



```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/924,345
FILING DATE: 04-SEP-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/199,172
FILING DATE: 09-AUG-1994
APPLICATION NUMBER: FR 92 07930
FILING DATE: 26-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: SARRO, THOMAS P.
REGISTRATION NUMBER: 19396
REFERENCE/DOCKET NUMBER: XI
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 920-7200
TELEFAX: (703) 892-8428
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 380 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-924-345-2

```

```

alignment_scores:
Quality: 55.50 Length: 28
Ratio: 3.265 Gaps: 1
Percent Similarity: 60.714 Percent Identity: 50.000

```

alignment_block:

```

US-09-712-768-7/rev x US-08-924-345-2 ..
Align seg 1/1 to: US-08-924-345-2 from: 1 to: 380

```

```

70 CGAGCTGCTGTTTCCGACATCGCAGCGC..... 41
||||| :||||| |||:||||
321 ArgAgaGagLeuLeuProCySerArgArgValIlyIlySProArgThrcy 337
40 .GCAGATCGGATCAGGACAGCAGCAGGAGAGA 8
||||:||||| |||:|||||
337 sAlaAlaCysGlySerGlyThrcySAlaGlyArg 348

```

seq_name: /cgn2.6/ptodata/2/1aa/5B_COMB.pep:US-09-025-819-2

seq_documentation_block:

```

Sequence 2, Application US/09025819
Patent No. 6225097
GENERAL INFORMATION:
APPLICANT: Odata, Shusei
APPLICANT: Nishino, Tokuzo
APPLICANT: Koyama, Tanetoshi
APPLICANT: Sato, Yoshihiro
TITLE OF INVENTION: DECAPRENYL DIPHOSPHATE SYNTHETASE GENE
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: KENYON & KENYON
STREET: 1500 K Street, N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/025,819
FILING DATE: 19-FEB-1998

```

```

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 251675
FILING DATE: 17-SEP-1997
ATTORNEY/AGENT INFORMATION:
NAME: Khalilian, Hourl
REGISTRATION NUMBER: 39,546
REFERENCE/DOCKET NUMBER: 10235/2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-220-4200
TELEFAX: 202-220-4201
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 331 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-025-819-2

```

```

alignment_scores:
Quality: 54.00 Length: 29
Ratio: 2.700 Gaps: 2
Percent Similarity: 68.966 Percent Identity: 51.724

```

alignment_block:

```

US-09-712-768-7/rev x US-09-025-819-2 ..
Align seg 1/1 to: US-09-025-819-2 from: 1 to: 331

```

```

85 CGGCATCTCGAAGCAGCGAGTGT...GTTTCGACATCGCAGCGCGC 39
||||| |||:||||| ||| ||| |||
252 ArgProArgArgArgGlyArgAlaArgLeuLeuGlyThrcHisHisArgPr 268
38 AGGA...TGGCGATCAGCAGCAGCAGCAGGAGATCG 5
||| :||| ||| ||| ||| |||
268 oGlyProAlaGlyArgGlyArgProGlyHisArgAla 280

```

seq_name: /cgn2.6/ptodata/2/1aa/backfill.es1.pep:5183745-2

seq_documentation_block:

```

Patent No. 5183745
APPLICANT: DANCHIN, ANTOINE;GLASER, PHILIPPE;KRIN, EVELYN;
BARZU, OCTAVIEN;LADANT, DANIEL;ULLMAN, AGNES
TITLE OF INVENTION: ADENYL CYCLASE DERIVATIVES AND THEIR
BIOLOGICAL USES
NUMBER OF SEQUENCES: 13
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/426,541
FILING DATE: 25-OCT-1989
SEQ ID NO: 2
LENGTH: 1489
5183745-2

```

```

alignment_scores:
Quality: 53.00 Length: 18
Ratio: 4.077 Gaps: 0
Percent Similarity: 72.222 Percent Identity: 50.000

```

alignment_block:

```

US-09-712-768-7/rev x 5183745-2 ..
Align seg 1/1 to: 5183745-2 from: 1 to: 1489

```

```

83 GCCATCTCGAAGCAGCGAGTGTTCGACATCGCAGCGCGCAGAGAT 34
||||| :||||| ||| :|||
1413 AlaMetArgArgProThrcCysAlaAlaArgValAlaProThrcCysLe 1429
33 GCGG 30
|||
1429 uArg 1430

```

```
seq_name: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:US-08-996-139-2
seq_documentation_block:
; Sequence 2, Application US/08996139
; Patent No. 6017729
; GENERAL INFORMATION:
; APPLICANT: Anderson, Dirk M.
; APPLICANT: Galibert, Laurent
; APPLICANT: Maraskovsky, Eugene
; TITLE OF INVENTION: Receptor Activator of NF-kappaB
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation, Law Department
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Power Macintosh
; OPERATING SYSTEM: Apple Operating System 7.5.5
; SOFTWARE: Microsoft Word for Power Macintosh 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/996,139
; FILING DATE: 22 DECEMBER 1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 60/064,671
; FILING DATE: 14 OCTOBER 1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/813,509
; FILING DATE: 07 MARCH 1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/772,330
; FILING DATE: 23 DECEMBER 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia Anne
; REGISTRATION NUMBER: 34,693
; REFERENCE/DOCKET NUMBER: 2851-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)587-0430
; TELEFAX: (206)233-0644
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 591 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-996-139-2

alignment_scores:
Quality: 51.50 Length: 19
Ratio: 3.219 Gaps: 1
Percent Similarity: 84.211 Percent Identity: 63.158

alignment_block:
US-09-712-768-7/rev x US-08-996-139-2
Align seg 1/1 to: US-08-996-139-2 from: 1 to: 591
58 TTGCGGACATCGACCGCGAGATCGGATCAGGC...AGACCGACGAG 12
|||||:||||| |||||||:||||| |||:|
476 LeupProSerSerAlaArgAlaGlyAlaGlySerGlySerSerProGlyI 492
11 AAGATCG 5
|:|:|:|
492 yGlnSer 494

seq_name: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:US-08-995-659-2
seq_documentation_block:
```

```
; Sequence 2, Application US/08995659
; Patent No. 6242213
; GENERAL INFORMATION:
; APPLICANT: Anderson, Dirk M.
; APPLICANT: Galibert, Laurent
; APPLICANT: Maraskovsky, Eugene
; TITLE OF INVENTION: Ligand for Receptor Activator of NF-kappaB
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation, Law Department
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Power Macintosh
; OPERATING SYSTEM: Apple Operating System 7.5.5
; SOFTWARE: Microsoft Word for Power Macintosh 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/995,659
; FILING DATE: 22 DECEMBER 1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 60/064,671
; FILING DATE: 14 OCTOBER 1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/813,509
; FILING DATE: 07 MARCH 1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/772,330
; FILING DATE: 23 DECEMBER 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia Anne
; REGISTRATION NUMBER: 34,693
; REFERENCE/DOCKET NUMBER: 2852-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)587-0430
; TELEFAX: (206)233-0644
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 591 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-995-659-2

alignment_scores:
Quality: 51.50 Length: 19
Ratio: 3.219 Gaps: 1
Percent Similarity: 84.211 Percent Identity: 63.158

alignment_block:
US-09-712-768-7/rev x US-08-995-659-2
Align seg 1/1 to: US-08-995-659-2 from: 1 to: 591
58 TTGCGGACATCGACCGCGAGATCGGATCAGGC...AGACCGACGAG 12
|||||:||||| |||||||:||||| |||:|
476 LeupProSerSerAlaArgAlaGlyAlaGlySerGlySerSerProGlyI 492
11 AAGATCG 5
|:|:|:|
492 yGlnSer 494

seq_name: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:US-09-215-649A-2
seq_documentation_block:
```

```

Sequence 2, Application US/09215649A
Patent No. 6271349
GENERAL INFORMATION:
APPLICANT: Anderson, Dirk M.
Galibert, Laurent
Maraskovsky, Eugene
TITLE OF INVENTION: Receptor Activator of NF-kappaB
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation, Law Department
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: Apple Power Macintosh
OPERATING SYSTEM: Apple Operating System 7.5.5
SOFTWARE: Microsoft Word for Power Macintosh 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/215,649A
FILING DATE: 17-Dec-1998
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/996,139
FILING DATE: <unknown>
APPLICATION NUMBER: USSN 08/813,509
FILING DATE: 07 MARCH 1997
APPLICATION NUMBER: USSN 08/772,330
FILING DATE: 23 DECEMBER 1996
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne
REGISTRATION NUMBER: 34,693
REFERENCE/DOCKET NUMBER: 2851-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)233-0644
TELEFAX: (206)587-0430
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 591 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-215-649A-2

alignment_scores:
Quality: 51.50 Length: 19
Ratio: 3.219 Gaps: 1
Percent Similarity: 84.211 Percent Identity: 63.158

alignment_block:
US-09-712-768-7/rev x US-09-215-649A-2 ..
Align seg 1/1 to: US-09-215-649A-2 from: 1 to: 591
58 TTGCCGACATCGCACCGCGAGATGCGGATCAGGC...AGACGAGCAGG 12
|||||:||||| |||||||:||||| |||:|||||
476 LeuProSerSerAlaArgAlaGlyAlaGlySerGlySerSerProGlyGI 492
11 AAGATCG 5
|:|:|
492 yGlnSer 494

seq_name: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:US-08-996-139-6
seq_documentation_block:
Sequence 6, Application US/08996139
Patent No. 6017729
GENERAL INFORMATION:
APPLICANT: Anderson, Dirk M.

```

```

APPLICANT: Galibert, Laurent
Maraskovsky, Eugene
TITLE OF INVENTION: Receptor Activator of NF-kappaB
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation, Law Department
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: Apple Power Macintosh
OPERATING SYSTEM: Apple Operating System 7.5.5
SOFTWARE: Microsoft Word for Power Macintosh 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/996,139
FILING DATE: 22 DECEMBER 1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 60/064,671
FILING DATE: 14 OCTOBER 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/813,509
FILING DATE: 07 MARCH 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/772,330
FILING DATE: 23 DECEMBER 1996
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne
REGISTRATION NUMBER: 34,693
REFERENCE/DOCKET NUMBER: 2851-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)233-0644
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 616 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-996-139-6

alignment_scores:
Quality: 51.50 Length: 19
Ratio: 3.219 Gaps: 1
Percent Similarity: 84.211 Percent Identity: 63.158

alignment_block:
US-09-712-768-7/rev x US-08-996-139-6 ..
Align seg 1/1 to: US-08-996-139-6 from: 1 to: 616
58 TTGCCGACATCGCACCGCGAGATGCGGATCAGGC...AGACGAGCAGG 12
|||||:||||| |||||||:||||| |||:|||||
501 LeuProSerSerAlaArgAlaGlyAlaGlySerGlySerSerProGlyGI 517
11 AAGATCG 5
|:|:|
517 yGlnSer 519

seq_name: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:US-08-996-659-6
seq_documentation_block:
Sequence 6, Application US/08996559
Patent No. 6242213
GENERAL INFORMATION:
APPLICANT: Anderson, Dirk M.
APPLICANT: Galibert, Laurent
APPLICANT: Maraskovsky, Eugene
TITLE OF INVENTION: Ligand for Receptor Activator of NF-kappaB

```

```

;
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation, Law Department
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Power Macintosh
; OPERATING SYSTEM: Apple Operating System 7.5.5
; SOFTWARE: Microsoft Word for Power Macintosh 6.0.1
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/995,659
; FILING DATE: 22 DECEMBER 1997
; CLASSIFICATION:
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USN 60/064,671
; FILING DATE: 14 OCTOBER 1997
; CLASSIFICATION:
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USN 08/813,509
; FILING DATE: 07 MARCH 1997
; CLASSIFICATION:
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USN 08/772,330
; FILING DATE: 23 DECEMBER 1996
; CLASSIFICATION:
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia Anne
; REGISTRATION NUMBER: 34,693
; REFERENCE/DOCKET NUMBER: 2852-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)587-0430
; TELEFAX: (206)233-0644
;
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 616 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-995-659-6

```

alignment_scores:

Quality:	51.50	Length:	19
Ratio:	3.219	Gaps:	1
Percent Similarity:	84.211	Percent Identity:	63.158

alignment_block:

US-09-712-768-7/rev x US-08-995-659-6 ..

Align seg 1/1 to: US-08-995-659-6 from: 1 to: 616

```

58 TTGCCGACATCGACCGCGAGATCGGATCAGGC...AGACGCGCAGG 12
|||||:||||| |||||||:|||||:||||| |||:|
501 LeuProSerSerAlaArgAlaGlyAlaGlySerGlySerProGlyGI 517
11 AGATCG 5
|:|:|
517 yGlnSer 519

```

seq_name: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:US-09-215-649A-6

seq_documentation_block:

Sequence 6, Application US/09215649A

PATENT No. 6271349

GENERAL INFORMATION:

APPLICANT: Anderson, Dirk M.
Galbert, Laurent
Maraskovsky, Eugene
TITLE OF INVENTION: Receptor Activator of NF-kappaB

```

;
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation, Law Department
; STREET: 51 University Street
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98101
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Power Macintosh
; OPERATING SYSTEM: Apple Operating System 7.5.5
; SOFTWARE: Microsoft Word for Power Macintosh 6.0.1
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/215,649A
; FILING DATE: 17-Dec-1998
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/996,139
; FILING DATE: <Unknown>
; APPLICATION NUMBER: USN 08/813,509
; FILING DATE: 07 MARCH 1997
; APPLICATION NUMBER: USN 08/772,330
; FILING DATE: 23 DECEMBER 1996
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia Anne
; REGISTRATION NUMBER: 34,693
; REFERENCE/DOCKET NUMBER: 2851-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)587-0430
; TELEFAX: (206)233-0644
;
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 616 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
; US-09-215-649A-6

```

alignment_scores:

Quality:	51.50	Length:	19
Ratio: <td>3.219<td>Gaps:<td>1</td></td></td>	3.219 <td>Gaps:<td>1</td></td>	Gaps: <td>1</td>	1
Percent Similarity: <td>84.211<td>Percent Identity:<td>63.158</td></td></td>	84.211 <td>Percent Identity:<td>63.158</td></td>	Percent Identity: <td>63.158</td>	63.158

alignment_block:

US-09-712-768-7/rev x US-09-215-649A-6 ..

Align seg 1/1 to: US-09-215-649A-6 from: 1 to: 616

```

58 TTGCCGACATCGACCGCGAGATCGGATCAGGC...AGACGCGCAGG 12
|||||:||||| |||||||:|||||:||||| |||:|
501 LeuProSerSerAlaArgAlaGlyAlaGlySerGlySerProGlyGI 517
11 AGATCG 5
|:|:|
517 yGlnSer 519

```

seq_name: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:US-08-726-306A-173

seq_documentation_block:

Sequence 173, Application US/08726306A

PATENT No. 5958684

GENERAL INFORMATION:

APPLICANT: van Leeuwen, Frederik Willem
Burdach, Johannes Peter Henri
APPLICANT: Grosveld, Franklin G.
TITLE OF INVENTION: DIAGNOSIS METHOD AND REAGENTS
NUMBER OF SEQUENCES: 189
CORRESPONDENCE ADDRESS:
BANNER & WITCOFF, LTD.
STREET: 1 Financial Center

CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02111
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 6.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/726,306A
FILING DATE: 02-Oct-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 95/20080.4
FILING DATE: 02-Oct-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/009,832
FILING DATE: 01-Jan-1996
ATTORNEY/AGENT INFORMATION:
NAME: Williams, Ph.D., Kathleen M.
REGISTRATION NUMBER: 34,380
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 345-9100
TELEFAX: (617) 345-9111
INFORMATION FOR SEQ ID NO: 173:
SEQUENCE CHARACTERISTICS:
LENGTH: 80 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-726-306A-173

alignment_scores:
Quality: 50.00 Length: 15
Ratio: 4.167 Gaps: 0
Percent Similarity: 80.000 Percent Identity: 66.667

alignment_block:
US-09-712-768-7 x US-08-726-306A-173 ..

Align seg 1/1 to: US-08-726-306A-173 from: 1 to: 80

8 TCTTCCTGCTGCTGCTGATCCGATCCGCGGTGCGATGT 52
:::||||| ||||| ||||| |||||:::|||||
39 AAtSerCysThrSerAlaArgSerAlaArgCysAlaAlaArgCys 53

seq_name: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:US-08-296-014A-4

seq_documentation_block:
Sequence 4, Application US/08296014A
Patent No. 5716834
GENERAL INFORMATION:
APPLICANT: Ding, Jeak Ling
APPLICANT: Ho, Bow
TITLE OF INVENTION: The Cloned Factor C cDNA of the
TITLE OF INVENTION: Singapore Horseshoe Crab, Carcinoscopus
TITLE OF INVENTION: rotundicauda and Purification of Factor C Proenzyme
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch
STREET: 8110 Gatehouse Road, Suite 500 East
CITY: Falls Church
STATE: Virginia
COUNTRY: USA
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/296,014A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murphy, Jr., Gerald M.
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 1781-105P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1019 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-296-014A-4

alignment_scores:
Quality: 50.00 Length: 35
Ratio: 2.941 Gaps: 1
Percent Similarity: 48.571 Percent Identity: 31.429

alignment_block:
US-09-712-768-7 x US-08-296-014A-4 ..

Align seg 1/1 to: US-08-296-014A-4 from: 1 to: 1019

17 TGTCTGCTGATCCGATCCTGCGCGTGCATGTGCGCA..... 58
|||||::: ||||||| ||| :::
741 TtpSerGlyArgProAlaSerCysIleProValCysGlyArgSerAsp 757
59AACAGCACGTCGCTTCGAGAT 80
757 rProArgSerProHeIleTrpAsnGlyAsnSerThrGluIleGlyLnt 774
81 GGCCG 85
|||||
774 rPro 775

seq_name: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:US-08-596-405-4

seq_documentation_block:
Sequence 4, Application US/08596405
Patent No. 5858706
GENERAL INFORMATION:
APPLICANT: Ding, Jeak Ling
APPLICANT: Ho, Bow
TITLE OF INVENTION: The Cloned Factor C cDNA of the
TITLE OF INVENTION: Singapore Horseshoe Crab, Carcinoscopus
TITLE OF INVENTION: rotundicauda and Purification of Factor C Proenzyme
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch
STREET: 8110 Gatehouse Road, Suite 500 East
CITY: Falls Church
STATE: Virginia
COUNTRY: USA
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/596,405
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murphy, Jr., Gerald M.

REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 1781-105P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1019 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-596-405-4

alignment_scores:
Quality: 50.00 Length: 35
Ratio: 2.941 Gaps: 1
Percent Similarity: 48.571 Percent Identity: 31.429

alignment_block:
US-09-712-768-7 x US-08-596-405-4 ..

Align seg 1/1 to: US-08-596-405-4 from: 1 to: 1019

17 TGGTCTGCTGATCCGATCCTGCGGCTGCGATGCGCA..... 58
|||||:|:| |||||
741 TrpSerGlyArgProAlaSerCysIleProValCysGlyArgSerAspse 757
59AACAGCAGCTCGGTTTCAGAT 80
|||||:|:| |||||
757 rProArgSerProPheIleTrpAsnGlyAsnSerThrGluIleGlyInt 774
81 GGCGG 85
|||||
774 rPPro 775

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2002, 01:56:27 ; Search time 31.32 seconds
(without alignments)
89.872 Million cell updates/sec

Title: US-09-712-768-6
Perfect score: 203
Sequence: 1 IVHGDRKTAIGLAIGLWFTLCQAVEYEVHTE 38

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 522463 seqs, 74073290 residues
Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: A.Geneseq_1101.*
2: /SIDS2/gcgdata/geneseq/geneseq/AA1980.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq/AA1981.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq/AA1982.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq/AA1983.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq/AA1984.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq/AA1985.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq/AA1986.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq/AA1987.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq/AA1988.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq/AA1989.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq/AA1990.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq/AA1991.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq/AA1992.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq/AA1993.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq/AA1994.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq/AA1995.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq/AA1996.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq/AA1997.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq/AA1998.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq/AA1999.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq/AA2000.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	203	100.0	38	22	AA97752
2	62	30.5	93	22	AAB63840
3	62	30.5	133	22	AAB29910
4	62	30.5	203	22	AAB75604
5	62	30.5	455	20	AA935610
6	60	29.6	133	21	AAB29909
7	60	29.6	172	21	AAB36196
8	60	29.6	203	22	AAB75603
9	58	28.6	612	22	AAB93232
10	58	28.6	747	20	AA941758
11	58	28.6	747	21	AAB44314

12	58	28.6	747	22	AA012384	Human PRO1012 poly
13	58	28.6	768	22	AAB92743	Human protein sequ
14	58	28.6	793	22	AAB93204	Human protein sequ
15	58	28.6	793	22	AAB95309	Human protein sequ
16	58	28.6	793	22	AAB71913	Human ISOM-5. Hom
17	55	27.1	78	22	AAB75551	Human secreted pro
18	55	27.1	93	22	AAG75934	Human colon cancer
19	55	27.1	847	22	AAG92548	C glutamylam prote
20	54.5	26.8	131	22	AAB74197	OmPA-J domain fusi
21	54.5	26.8	131	22	AAB70767	Expression plasmid
22	54.5	26.8	131	22	AA972018	E. coli OmPA-J dom
23	54.5	26.8	399	21	AAB11395	E. coli expression
24	54.5	26.8	399	22	AAB74196	OmPA-DnaJ fusion p
25	54.5	26.8	399	22	AAB70766	Expression plasmid
26	54.5	26.8	399	22	AA972017	E. coli OmPA-DnaJ
27	54	26.6	48	22	AA950908	C glutamylam prote
28	53.5	26.4	131	21	AAB11396	E. coli expression
29	53.5	26.4	326	21	AA946264	Arabidopsis thalia
30	53.5	26.4	336	21	AA946263	Arabidopsis thalia
31	53.5	26.4	1147	21	AA930843	Arabidopsis thalia
32	53.5	26.4	1215	21	AA930842	Arabidopsis thalia
33	53.5	26.4	1355	21	AA930841	Arabidopsis thalia
34	53	26.1	117	20	AA900366	Fragment of human
35	53	26.1	216	22	AA900872	Human cancer relat
36	53	26.1	259	20	AA950556	Human SBHNC50 pro
37	53	26.1	276	19	AA970321	Secreted protein C
38	53	26.1	488	21	AA96758	A. terreus ORF16 M
39	53	26.1	1816	21	AA95440	Caenorhabditis ele
40	51.5	25.4	192	20	AA931662	Staphylococcus aur
41	51	25.1	376	22	AA003811	G protein-coupled
42	50.5	24.9	265	20	AA909510	Human M6b1 protein
43	50.5	24.9	870	20	AA935233	C. pneumoniae prote
44	50	24.6	205	22	AA92148	C glutamylam prote
45	50	24.6	217	22	AAB79574	Corynebacterium g1

ALIGNMENTS

RESULT 1
ID AAY97752 standard; Protein: 38 AA.
AC AAY97752;

06-AUG-2001 (first entry)

G. oxydans cytochrome C oxidase (COIII) protein sequence fragment.

Cytochrome C oxidase; COI; COII; COIII; COI complex; D-sorbitol;
oxidative fermentation; electron transfer; respiratory chain; L-sorbose;
2-keto-L-gluconic acid; 2KGa production; aldehyde production;
carboxylic acid production; ketone production.

Glucanobacter oxydans.

EP103603-A2.

30-MAY-2001.

14-NOV-2000; 2000EP-0124785.

17-NOV-1999; 99EP-0122842.

(HOFF) HOFFMANN LA ROCHE & CO AG F.

Asakura A, Hoshino T, Shinjoh M;

WPI: 2001-357953/38.

N-PSDB; AAA91491.

New cytochrome c oxidase complex having cytochrome c oxidase activity
from Glucanobacter oxydans DSM 4025, useful in mediating electron

PT transfer in respiratory chain or producing 2-keto-L-gulononic acid from
 PT L-sorbose or D-sorbitol
 XX
 PS Claim 6; Page 25; 42pp; English.

CC This sequence represents a fragment of the Gluconobacter oxydans
 CC cytochrome c oxidase, COI11. The invention relates to a cytochrome c
 CC oxidase COI complex. The COI complex is useful in improving oxidative
 CC fermentation and is an essential component mediating electron transfer
 CC in the respiratory chain. The recombinant microorganism and the
 CC cytochrome c oxidase may be used in the genetic preparation of a
 CC recombinant COI complex and in the production of 2-keto-L-gulononic acid
 CC (2KGA) from L-sorbose or D-sorbitol in a culture medium. The COI is also
 CC useful as a terminal oxidase, oxidizing cytochrome c, an electron
 CC acceptor from an enzyme belonging to dehydrogenase for the production of
 CC aldehydes, carboxylic acids and ketones from alcohols and aldehydes,
 CC especially the production of 2KGA from L-sorbose or D-sorbitol.

XX Sequence 38 AA;

Query Match 100.0%; Score 203; DB 22; Length 38;
 Best Local Similarity 100.0%; Pred. No. 2.1e-23;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IVHGDRKRTAIGLAIAIGLWIFLCAQYEEYVEIWHTE 38
 Db 1 lvhgdrkrtalgaiaiglwiflcaqeyeyelwhte 38

RESULT 2
 AAB63840 AAB63840 standard; Protein; 93 AA.

XX AAB63840;

AC AAB63840;

DT 26-MAR-2001 (first entry)

DE Human prostate cancer associated antigen protein sequence SEQ ID NO:1202.

XX Human: breast cancer; gastric cancer; prostate cancer; diagnosis;

KW cancer associated antigen; cytosolic; cancer vaccine.

OS Homo sapiens.

XX WO200073801-A2.

XX 07-DEC-2000.

XX 26-MAY-2000; 2000MO-US14749.

XX 28-MAY-1999; 99US-0136526.

PR 10-SEP-1999; 99US-0153454.

XX (LUDW-) LUDWIG INST CANCER RES.

XX Obata Y;

XX WPI; 2001-025274/03.

XX Nucleic acids encoding breast, gastric and prostate cancer associated
 PT antigen precursors, useful for diagnosing and treating a condition
 PT characterized by expression of an abnormal amount of a protein, e.g.
 PT cancer

XX Example 1; Page 741; 799pp; English.

XX AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
 CC represent nucleotide sequences encoding human breast, gastric and
 CC prostate cancer associated antigen precursors (CAAP) respectively.
 CC AAB63332 to AAB63467, AAB63468 to AAB63721 and AAB63722 to AAB63970
 CC represent human breast, gastric and prostate CAAP protein sequence
 CC respectively. CAAPs have cytostatic activity and can be used in the

CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
 CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
 CC condition characterised by expression of an abnormal amount of a protein,
 CC e.g. cancer.

XX Sequence 93 AA;

Query Match 30.5%; Score 62; DB 22; Length 93;
 Best Local Similarity 42.4%; Pred. No. 0.088;
 Matches 14; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

OY 1 IVHGDRKRTAIGLAIAIGLWIFLCAQYEEYVEI 33
 Db 5 llnnngqilqallitlllglyftllqaseyfe 37

RESULT 3
 AAB29910 AAB29910 standard; Protein; 133 AA.

XX AAB29910;

DT 09-FEB-2001 (first entry)

DE Human secreted protein BLAST search protein SEQ ID NO: 168.

KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; antitumor;
 KW vulnereary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.

XX Homo sapiens.

XX WO200061779-A1.

XX 19-OCT-2000.

XX 06-APR-2000; 2000MO-US09068.

XX 09-APR-1999; 99US-0128699.

XX 20-JAN-2000; 2000US-0177050.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM, Komatsoulis G;

XX WPI; 2000-647424/52.

XX Isolated nucleic acid molecule encoding a human secreted protein is
 PT used in preventing, treating or ameliorating a medical condition -
 PT disclosure; Page 488-489; 495pp; English.

XX The invention relates to the isolation of genes AAF63410-C63458 encoding
 CC the human secreted proteins AAB29802-B29850. This sequence represents a
 CC fragment of the protein encoded by the gene given in the descriptor
 CC line. The sequence is used as a query sequence for doing BLASTX searches
 CC to determine homologous sequence to the protein. The genes and proteins
 CC are useful for preventing, ameliorating or treating medical conditions,
 CC e.g. by protein or gene therapy. The genes are isolated from a range of
 CC human tissues disclosed in the specification. The nucleic acids,
 CC proteins, antibodies and (ant)agonists are useful in the diagnosis,
 CC treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer,
 CC and other cancers of the adrenal gland, bone, bone marrow, breast,
 CC gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders
 CC e.g. Addison's disease, allergies, autoimmune haemolytic anaemia,
 CC autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple
 CC sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
 CC cardiovascular disorders such as myocardial ischaemias; (d) wound
 CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and
 CC (f) infectious diseases such as viral, bacterial, fungal and parasitic

CC infections.
XX
SQ Sequence 133 AA;

Query Match 30.5%; Score 62; DB 21; Length 133;
Best Local Similarity 42.4%; Pred. No. 0.13;
Matches 14; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

OY 1 IVHGDRKRTAIGLAIGLWFTLCQAVEYFE 33
Db 23 Ilenmrnglqallllllllyflilgaseyfe 55

RESULT 4
AAB75604
ID AAB75604 standard; Protein; 203 AA.

AC AAB75604;

DT 06-APR-2001 (first entry)

DE Human secreted protein sequence encoded by gene 46 spq ID NO:158.

XX Human: secreted protein; immunosuppressive; antiarthritic; antirheumatic;
XX antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;
XX neurotropic; neuroprotective; antibacterial; virucide; fungicide;
XX ophthalmological; vulnery; autoimmune disease; cardiovascular disorder;
XX hyperproliferative disorders; cerebrovascular disorder; wound healing;
XX nervous system disorder; ocular disorder; skin ageing; chemotaxis;
XX food additive.

OS Homo sapiens.

PN W0200077026-A1.

PD 21-DEC-2000.

XX 01-JUN-2000; 2000WO-US14973.

XX 11-JUN-1999; 99US-0138630.

XX (HUMA-) HUMAN GENOME SCI INC.

XX (ROSE/) ROSEN C A.

XX Rosen CA, Ruben SM, Komatsoulis GA;

XX WPI: 2001-071258/08.

XX N-PSDB; AAF64221.

XX Nucleic acid molecules encoding human secreted proteins, used in

XX preventing, treating or ameliorating a disorder, e.g. Alzheimer's and

XX Parkinson's diseases and cancers -

XX Disclosure: Page 94; 542pp; English.

XX Human secreted proteins AAB75506 - AAB75554 are encoded by polynucleotide

XX sequences AAF64176 - AAF64224. The specification includes amino acid

XX sequences AAB75555 - AAB75606 which represent fragments of the human

XX secreted proteins, and protein sequences with which they share homology.

XX The proteins and polynucleotides, their agonists and antagonists have

XX activities dependent on the tissues and cells in which they are

XX expressed, examples of these activities include: immunosuppressive;

XX antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiant;

XX vasotropic; cerebroprotective; neurotropic; neuroprotective; antibacterial;

XX virucide; fungicide; ophthalmological; and vulnery. The proteins, AAB75506

XX polynucleotides, agonists and antagonists can be used to treat or detect

XX or diagnose various diseases and disorders including: autoimmune

XX diseases e.g. rheumatoid arthritis, hyperproliferative disorders

XX e.g. neoplasms of the breast or liver, cardiovascular disorders

XX e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,

XX angiogenesis, nervous system disorders e.g. Alzheimer's disease,

XX infections caused by bacteria, viruses and fungi and ocular disorders

CC e.g. corneal infection. The polypeptides can also be used to aid wound
CC healing and epithelial cell proliferation, to prevent skin ageing due to
CC sunburn, to maintain organs before transplantation, for supporting cell
CC culture of primary tissues, to regenerate tissues and in chemotaxis. The
CC polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities. Included in the invention are
CC polynucleotide sequences AAF64167 - AAF64175 and peptide AAB75505 which
CC are used in the isolation, identification and characterisation of the
CC proteins of the invention.

XX Sequence 203 AA;

Query Match 30.5%; Score 62; DB 22; Length 203;
Best Local Similarity 42.4%; Pred. No. 0.22;
Matches 14; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

OY 1 IVHGDRKRTAIGLAIGLWFTLCQAVEYFE 33
Db 93 Ilenmrnglqallllllllyflilgaseyfe 125

RESULT 5
AAV35610
ID AAV35610 standard; Protein; 455 AA.

AC AAV35610;

DT 13-SEP-1999 (first entry)

DE Chlamydia pneumoniae lipoprotein sequence.

XX Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
XX sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
XX vaccine; neutralising epitope.

XX Chlamydia pneumoniae.

XX W09927105-A2.

XX 03-JUN-1999.

XX 20-NOV-1998; 98WO-1B01890.

XX 04-NOV-1998; 98US-0107078.

XX 21-NOV-1997; 97FR-0014673.

XX (GEST) GENSET.

XX Griffiths R;

XX WPI: 1999-357842/30.

XX Genome sequence of Chlamydia pneumoniae

XX Page 1338-1339; Disclosure: 1912pp; English.

XX AAV34584-Y35879 represent the proteins encoded by all the open reading

XX frames in the complete genome (see AAV31990) of Chlamydia pneumoniae.

XX C. pneumoniae causes respiratory disease such as pneumonia and

XX bronchitis and is thought to be a contributing factor in heart

XX disease, sarcoidosis, sinusitis, purulent otitis media, erythema

XX nodosum or pharyngitis. The polypeptides encoded by the open reading

XX frames of the C. pneumoniae genome (see AAV34584-Y35879) can be used in

XX immunogenic compositions as vaccines. Vectors containing C. pneumoniae

XX nucleotide sequences can also be used as immunogenic compositions,

XX especially where the vector directs the expression of a neutralising

XX epitope of C. pneumoniae.

SQ Sequence 455 AA;

Query Match 30.5%; Score 62; DB 20; Length 455;

Best Local Similarity 50.0%; Pred. No. 0.56;
Matches 9; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 21 WFTLCAVEYEVITE 38
Db 35 WYLCQSYEHKKLVFPE 52

RESULT 6

AAB29909

ID AAB29909 standard; Protein; 133 AA.

AC AAB29909;

DE 09-FEB-2001 (first entry)

Human secreted protein BLAST search protein SEQ ID NO: 167.

KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antinflammatory; antitumor;
KW vulnerrary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein.

OS Homo sapiens.

XX WO200061779-A1.

PD 19-OCT-2000.

PF 06-APR-2000; 2000WO-US09068.

PR 09-APR-1999; 99US-0128699.

PR 20-JAN-2000; 2000US-0177050.

XX (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM, Komatsoulis G;

DR WPI: 2000-647424/62.

PT Isolated nucleic acid molecule encoding a human secreted protein is
PT used in preventing, treating or ameliorating a medical condition -
XX

PS Disclosure: Page 487; 495pp; English.

CC The invention relates to the isolation of genes AACC3410-C63458 encoding
CC the human secreted proteins AAB29802-B29850. This sequence represents a
CC peptide fragment homologous to the protein encoded by the gene given
CC in the descriptor line. The sequence is a search result from a BLASTX
CC homology search. The genes and proteins are useful for preventing,
CC ameliorating or treating medical conditions, e.g. by protein or gene
CC therapy. The genes are isolated from a range of human tissues disclosed
CC in the specification. The nucleic acids, proteins, antibodies and
CC (ant)agonists are useful in the diagnosis, treatment and prevention of:
CC (a) cancer, e.g. breast and ovarian cancer, and other cancers of the
CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
CC lung, or urogenital; (b) immune disorders e.g. Addison's disease,
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; (c) cardiovascular disorders such
CC as myocardial ischaemias; (d) wound healing; (e) neurological diseases
CC e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such
CC as viral, bacterial, fungal and parasitic infections.

XX Sequence 133 AA;

Query Match

Best Local Similarity 29.6%; Score 60; DB 21; Length 133;
Matches 14; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

QY 1 IVHGDRKTAIGLAIGWIFLCAVEY 33

Db 23 IImennrnmigalllillllyflllgaseyfe 55

RESULT 7

AAB36196

ID AAB36196 standard; Protein; 172 AA.

AC AAB36196;

DE 02-MAR-2001 (first entry)

Cowdria ruminatum 4hworfl polypeptide.

KW Cowdria ruminatum; MAP1; major antigenic protein 1; antirickettsial;
KW vaccine; gene therapy; Rickettsia; Cowdria; Anaplasma; map2; 1hworfl3;
KW 4hworfl; 18hworfl; 3gworfl3.

OS Cowdria ruminatum.

XX WO200065063-A2.

PD 02-NOV-2000.

PF 21-APR-2000; 2000WO-US10886.

PR 22-APR-1999; 99US-0130725.

XX (UYFL) UNIV FLORIDA.

PI Barbet AF, Bowie MV, Ganta RR, Burridge MJ, Mahan SM, McGuire TC;

PI Rurangirwa FR, Moreland AL, Simbi BH, Whitmore WW, Alleman AR;

DR N-PSDB; AAC68713.

PT New polynucleotides useful as DNA vaccines for conferring immunity to
PT rickettsial infection e.g. heartwater caused by Cowdria ruminatum,
PT comprises major antigenic protein 1 or 2 gene of rickettsial pathogens
XX

PS Claim 3; Page 58-59; 63pp; English.

CC The present sequence is given in a specification relating to nucleic
CC acid vaccines which may be used to protect animals or humans against
CC rickettsial diseases caused by a organisms of Rickettsia sp.,
CC Ehrlichia sp., Anaplasma sp. or Cowdria sp. The vaccine elicits an
CC immune response protective against the rickettsial pathogen. The
CC vaccine comprises the major antigenic protein 1 (MAP1) or major
CC antigenic protein 2 (MAP2) of rickettsial pathogens. The nucleic acid
CC enhancer-promoter. Cowdria ruminatum genes designated map 2, 1hworfl3,
CC 4hworfl, 18hworfl and 3gworfl may be used in therapeutic and diagnostic
CC applications. The polypeptides are useful for detecting antibodies
CC associated with infection by a rickettsial pathogen whilst the
CC polynucleotides may be used to detect the presence of rickettsial
CC nucleic acids.

XX Sequence 172 AA;

Query Match

Best Local Similarity 29.6%; Score 60; DB 21; Length 172;
Matches 14; Conservative 4; Mismatches 13; Indels 2; Gaps 1;

QY 1 IVHGDRK--KTATGALIAIGWIFLCAVEY 31

Db 63 ihwdrkyntnigkvtlgncllhaceidgy 95

RESULT 8

AAB75603

ID AAB75603 standard; Protein; 203 AA.

XX AAB75603;
 AC 06-APR-2001 (first entry)
 DT Gene 46 human secreted protein homologous amino acid sequence #157.
 DE
 XX Human: secreted protein; immunosuppressive; antiarthritic; antirheumatic;
 XX antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;
 XX neuroprotective; antibacterial; virocidic; fungicide; disorder;
 XX ophthalmological; vulnery; autoimmune disease; cardiovascular disorder;
 XX hyperproliferative disorders; cerebrovascular disorder; wound healing;
 XX nervous system disorder; ocular disorder; skin ageing; chemotaxis;
 XX food additive.
 XX Homo sapiens.
 OS WO200077026-A1.
 PN 21-DEC-2000.
 PD
 XX 01-JUN-2000; 2000WO-US14973.
 PF
 XX 11-JUN-1999; 99US-0138630.
 PR
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX (ROSE/) ROSEN C A.
 PA
 XX Rosen CA, Ruben SM, Komatsoulis GA;
 PI WPI: 2001-071258/08.
 DR Nucleic acid molecules encoding human secreted proteins, used in
 XX preventing, creating or ameliorating a disorder, e.g. Alzheimer's and
 XX Parkinson's diseases and cancers -
 XX
 XX Disclosure: Page 94; 542pp; English.
 PS
 XX Human secreted proteins AAB75506 - AAB75554 are encoded by polynucleotide
 CC sequences AAF64176 - AAF64224. The specification includes amino acid
 CC sequences AAB75555 - AAB75606 which represent fragments of the human
 CC secreted proteins, and protein sequences with which they share homology.
 CC The proteins and polynucleotides, their agonists and antagonists have
 CC activities dependent on the tissues and cells in which they are
 CC expressed, examples of these activities include: immunosuppressive;
 CC antirheumatic; antiproliferative; cytostatic; cardiant;
 CC vasotropic; cerebroprotective; neuroprotective; antibacterial;
 CC virocidic; fungicide; ophthalmological; and vulnery. The proteins,
 CC polynucleotides, agonists and antagonists can be used to treat or detect
 CC or diagnose various diseases and disorders including, autoimmune
 CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders
 CC e.g. neoplasms of the breast or liver, cardiovascular disorders
 CC e.g. cardiac arrest, cerebrovascular disorders e.g. Alzheimer's disease,
 CC angiogenesis, nervous system disorders e.g. Alzheimer's disease,
 CC infections caused by bacteria, viruses and fungi and ocular disorders
 CC e.g. corneal infection. The polypeptides can also be used to aid wound
 CC healing and epithelial cell proliferation, to prevent skin ageing due to
 CC sunburn, to maintain organs before transplantation, for supporting cell
 CC culture of primary tissues, to regenerate tissues and in chemotaxis. The
 CC polypeptides can also be used as a food additive or preservative to
 CC increase or decrease storage capabilities. Included in the invention are
 CC polynucleotide sequences AAF64167 - AAF64175 and peptide AAB75505 which
 CC are used in the isolation, identification and characterisation of the
 CC proteins of the invention.
 CC
 XX Sequence 203 AA;
 SQ

Query Match 29.6%; Score 60; DB 22; Length 203;
 Best Local Similarity 42.4%; Pred. No. 0.44;
 Matches 14; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

1 IVHGDRKKTATGTAIGLWITLQAYEYFE 33

DB 93 lmenrngmigaillitlllyftllqeseyfe 125
 RESULT 9
 ID AAB93232 standard: Protein: 612 AA.
 AC AAB93232;
 XX 26-JUN-2001 (first entry)
 DT Human protein sequence SEQ ID NO:12229.
 DE Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX Homo sapiens.
 OS EP1074617-A2.
 PN 07-FEB-2001.
 PD 28-JUL-2000; 2000EP-0116126.
 PF 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX (HELI-) HELIX RES INST.
 XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 DR WPI: 2001-318749/34.
 XX
 XX Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 PS
 XX Claim 8: SEQ ID 12229; 2537Pp + CD ROM; English.
 XX
 XX The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises at least 15 nucleotides and the combination
 CC of oligonucleotide comprises a 3'-end sequence is selected from those defined in
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 CC
 XX Sequence 612 AA;
 SQ

Query Match 28.6%; Score 58; DB 22; Length 612;
 Best Local Similarity 27.1%; Pred. No. 3.1;
 Matches 13; Conservative 3; Mismatches 8; Indels 24; Gaps 1;

OY 3 HDRKK-----TAIGLAIAIGLWIFLLC 26
 DB 27 hgdskeslvsfamghvstvtelwtgnfnvslqtaaglgwllfllc 74
 RESULT 10
 ID AAY41758 standard; Protein; 747 AA.
 AC AAY41758;
 DT 07-DEC-1999 (first entry)
 DE Human PRO1012 protein sequence.
 KW Human; PRO; EST; expressed sequence tag; PCR primer; hybridisation;
 KW probe; blood coagulation disorder; cancer; cellular adhesion disorder;
 XX secreted protein; transmembrane protein.
 OS Homo sapiens.
 PN WO9946281-A2.
 FD 16-SEP-1999.
 XX 08-MAR-1999; 99WO-US05028.
 XX 10-MAR-1998; 98US-0077450.
 PR 11-MAR-1998; 98US-0077632.
 PR 11-MAR-1998; 98US-0077641.
 PR 12-MAR-1998; 98US-0077649.
 PR 13-MAR-1998; 98US-0077791.
 PR 17-MAR-1998; 98US-0040220.
 PR 20-MAR-1998; 98US-0078886.
 PR 20-MAR-1998; 98US-0078910.
 PR 20-MAR-1998; 98US-0078936.
 PR 25-MAR-1998; 98US-0078939.
 PR 26-MAR-1998; 98US-0079294.
 PR 27-MAR-1998; 98US-0079656.
 PR 27-MAR-1998; 98US-0079664.
 PR 27-MAR-1998; 98US-0079689.
 PR 30-MAR-1998; 98US-0079786.
 PR 30-MAR-1998; 98US-0079920.
 PR 31-MAR-1998; 98US-0080105.
 PR 31-MAR-1998; 98US-0080107.
 PR 31-MAR-1998; 98US-0080165.
 PR 01-APR-1998; 98US-0080194.
 PR 01-APR-1998; 98US-0080327.
 PR 01-APR-1998; 98US-0080328.
 PR 01-APR-1998; 98US-0080333.
 PR 08-APR-1998; 98US-0080334.
 PR 08-APR-1998; 98US-0081049.
 PR 08-APR-1998; 98US-0081070.
 PR 09-APR-1998; 98US-0081071.
 PR 09-APR-1998; 98US-0081195.
 PR 09-APR-1998; 98US-0081203.
 PR 15-APR-1998; 98US-0081229.
 PR 15-APR-1998; 98US-0081817.
 PR 15-APR-1998; 98US-0081838.
 PR 15-APR-1998; 98US-0081952.
 PR 21-APR-1998; 98US-0081953.
 PR 21-APR-1998; 98US-0082568.
 PR 22-APR-1998; 98US-0082569.
 PR 22-APR-1998; 98US-0082700.
 PR 22-APR-1998; 98US-0082704.
 PR 23-APR-1998; 98US-0082804.
 PR 23-APR-1998; 98US-0082767.
 PR 23-APR-1998; 98US-0082796.

PR 27-APR-1998; 98US-0083336.
 PR 28-APR-1998; 98US-0083392.
 PR 29-APR-1998; 98US-0083392.
 PR 29-APR-1998; 98US-0083495.
 PR 29-APR-1998; 98US-0083496.
 PR 29-APR-1998; 98US-0083499.
 PR 29-APR-1998; 98US-0083500.
 PR 29-APR-1998; 98US-0083505.
 PR 29-APR-1998; 98US-0083545.
 PR 29-APR-1998; 98US-0083554.
 PR 29-APR-1998; 98US-0083558.
 PR 30-APR-1998; 98US-0083559.
 PR 05-MAY-1998; 98US-0083742.
 PR 06-MAY-1998; 98US-0084366.
 PR 06-MAY-1998; 98US-0084414.
 PR 07-MAY-1998; 98US-0084441.
 PR 07-MAY-1998; 98US-0084598.
 PR 07-MAY-1998; 98US-0084600.
 PR 07-MAY-1998; 98US-0084627.
 PR 07-MAY-1998; 98US-0084637.
 PR 07-MAY-1998; 98US-0084639.
 PR 07-MAY-1998; 98US-0084640.
 PR 07-MAY-1998; 98US-0084643.
 PR 13-MAY-1998; 98US-0085323.
 PR 13-MAY-1998; 98US-0085338.
 PR 15-MAY-1998; 98US-0085339.
 PR 15-MAY-1998; 98US-0085373.
 PR 15-MAY-1998; 98US-0085579.
 PR 15-MAY-1998; 98US-0085580.
 PR 15-MAY-1998; 98US-0085582.
 PR 15-MAY-1998; 98US-0085589.
 PR 15-MAY-1998; 98US-0085597.
 PR 15-MAY-1998; 98US-0085700.
 PR 18-MAY-1998; 98US-0085704.
 PR 22-MAY-1998; 98US-0086023.
 PR 22-MAY-1998; 98US-0086029.
 PR 22-MAY-1998; 98US-0086414.
 PR 22-MAY-1998; 98US-0086417.
 PR 28-MAY-1998; 98US-0086486.
 PR 28-MAY-1998; 98US-0087098.
 PR 28-MAY-1998; 98US-0087106.
 PR 30-JUL-1998; 98US-0087208.
 PR 11-SEP-1998; 98US-0094651.
 PR 11-SEP-1998; 98US-0100038.
 XX (GETH) GENENTECH INC.
 PA Wood WI, Goddard A, Gurney A, Yuan J, Baker KP, Chen J;
 PI WPI; 1999-551358/46.
 DR N-PSDB; AA234275.
 XX
 XX New secreted and transmembrane polypeptides and their polynucleotides,
 PT useful for treating blood coagulation disorders, cancers and cellular
 PT adhesion disorders -
 PS Claim 12; Fig 190; 530pp; English.
 XX
 XX The present invention describes secreted and transmembrane polypeptides
 CC and their polynucleotides. The nucleotide sequences are useful as
 CC sources of probes, primers, for chromosome mapping, and for generation
 CC of antisense sequences. They can also be used to treat a variety of diseases
 CC disorders, depending on their function. Diseases that may be treated
 CC include blood coagulation disorders, cancers and cellular adhesion
 CC disorders. They may also be used to raise antibodies. AA233891 to
 CC AA234338, and AAY1774 represent polynucleotide and
 CC polypeptide sequence given in the exemplification of the present
 CC invention.
 XX
 SQ Sequence 747 AA:

Query Match

28.6%; Score 58; DB 20; Length 747;

Best Local Similarity 27.1%; Pred. No. 4;
Matches 13; Conservative 3; Mismatches 8; Indels 24; Gaps 1;

OY 3 HGDRKK-----TAIGLAIAGLGMIFTLIC 26
| | | | | : | | | | |
Db 217 hgdrskesjvsfamghvstvtelwtgnfnvsigtatfaagigwiltfc 264

RESULT 11

ID AAB44314 standard; Protein; 747 AA.

AC AAB44314;

DT 08-FEB-2001 (first entry)

DE Human PRO1012 (UNQ495) protein sequence SEQ ID NO:459.

KW Human; secreted protein; transmembrane protein; PRO; EST; cytosolic;
expressed sequence tag; detection; cancer.

XX Homo sapiens.

PN WO200053756-A2.

PD 14-SEP-2000.

PF 18-FEB-2000; 2000WO-US04311.

PR 08-MAR-1999; 99WO-US05028.

PR 12-MAR-1999; 99US-0123957.

PR 29-MAR-1999; 99US-0126773.

PR 21-APR-1999; 99US-0130232.

PR 28-APR-1999; 99US-0131445.

PR 14-MAY-1999; 99US-0134287.

PR 23-JUN-1999; 99US-0141037.

PR 26-JUL-1999; 99US-0145698.

PR 29-OCT-1999; 99US-0162506.

PR 30-NOV-1999; 99WO-US28313.

PR 02-DEC-1999; 99WO-US28551.

PR 02-DEC-1999; 99WO-US28565.

PR 16-DEC-1999; 99WO-US30095.

PR 30-DEC-1999; 99WO-US31243.

PR 30-DEC-1999; 99WO-US31274.

PR 05-JAN-2000; 2000WO-US00219.

PR 06-JAN-2000; 2000WO-US00277.

PR 06-JAN-2000; 2000WO-US00376.

XX (GETH) GENENTECH INC.

PA Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;
PI Ferrara N, Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME;
PI Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ;
PI Kijavini IJ, Kuo SS, Nepier MA, Pan J, Paoni NF, Roy MA;
PI Shelton DL, Stewart TA, Tumas D, Williams PM, Wood WI;

XX WPI; 2000-611443/58.

DR N-PSDB; AAC78573.

XX Novel PRO polypeptides and polynucleotides used in detection methods,
PT to target bioactive molecules to specific cells, and to modulate
PI cellular activities -
XX Claim 12; Fig 190; 636pp; English.

CC The polypeptide pairs provide specific targeting of bioactive molecules
CC to cells. AAC78600 to AAC78987 represent PCR primers and probes used in
CC the isolation of the PRO polynucleotide sequences.

SO Sequence 747 AA;

Query Match 28.6%; Score 58; DB 21; Length 747;
Best Local Similarity 27.1%; Pred. No. 4;
Matches 13; Conservative 3; Mismatches 8; Indels 24; Gaps 1;

OY 3 HGDRKK-----TAIGLAIAGLGMIFTLIC 26
| | | | | : | | | | |
Db 217 hgdrskesjvsfamghvstvtelwtgnfnvsigtatfaagigwiltfc 264

RESULT 12

ID AAU12384 standard; Protein; 747 AA.

AC AAU12384;

DT 24-OCT-2001 (first entry)

DE Human PRO1012 polypeptide sequence.

KW Human secretory and transmembrane; PRO; mammalian; cancer; lung;
breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;
cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;
adipocyte; A-peptide; factor VIIA; gene therapy.

XX Homo sapiens.

PN WO200140466-A2.

PD 07-JUN-2001.

PF 01-DEC-2000; 2000WO-US32678.

PR 01-DEC-1999; 99WO-US28301.

PR 01-DEC-1999; 99WO-US28634.

PR 02-DEC-1999; 99WO-US28551.

PR 02-DEC-1999; 99WO-US28564.

PR 02-DEC-1999; 99WO-US28565.

PR 09-DEC-1999; 99US-0170262.

PR 16-DEC-1999; 99WO-US30095.

PR 20-DEC-1999; 99WO-US30911.

PR 20-DEC-1999; 99WO-US30999.

PR 30-DEC-1999; 99WO-US31243.

PR 06-JAN-2000; 2000WO-US00277.

PR 06-JAN-2000; 2000WO-US00376.

PR 11-FEB-2000; 2000WO-US03565.

PR 18-FEB-2000; 2000WO-US04341.

PR 18-FEB-2000; 2000WO-US04342.

PR 22-FEB-2000; 2000WO-US04414.

PR 24-FEB-2000; 2000WO-US04914.

PR 24-FEB-2000; 2000WO-US05004.

PR 01-MAR-2000; 2000WO-US05601.

PR 20-MAR-2000; 2000WO-US07377.

PR 21-MAR-2000; 2000WO-US07532.

PR 30-MAR-2000; 2000WO-US08439.

PR 17-MAY-2000; 2000WO-US13705.

PR 22-MAY-2000; 2000WO-US14042.

PR 30-MAY-2000; 2000WO-US14941.

PR 02-JUN-2000; 2000WO-US15264.

PR 10-NOV-2000; 2000WO-US30873.

XX (GETH) GENENTECH INC.

PA Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;

DR WPI: 2001-408281/43.
 DR N-PSDB: AAS21456.
 XX
 PT Isolated, secretory and transmembrane PRO polypeptide used to detect
 PT other PRO polypeptides, link bioactive molecules to cells expressing
 PT PRO polypeptides, and detect the presence of mammalian tumours e.g.
 PT lung, breast, prostate, cervical
 XX
 PS Claim 12; Fig 426; 813pp; English.
 XX
 CC AAU12172-AAU12446 represent novel human secretory and transmembrane
 CC PRO polypeptides. The PRO polypeptides are useful to detect other
 CC PRO polypeptides, to link bioactive molecules to cells expressing
 CC PRO polypeptides, to modulate biological activities of cells expressing
 CC breast, prostate, rectal, cervical or liver tumours by comparing PRO
 CC polypeptide expression in a cell sample to that in a control sample.
 CC Some of the 275 sequences are also useful to stimulate the release of
 CC tumor necrosis factor-alpha (TNF-alpha) from human blood, the
 CC proliferation or differentiation of chondrocytes, the proliferation or
 CC gene expression in pericyte cells, the release of proteoglycans from
 CC cartilage, the proliferation of inner ear utricular supporting cells or
 CC of T-lymphocytes, or the release of a cytokine from peripheral blood
 CC monocytes (PBMCs), or the proliferation of endothelial cells. Some of
 CC the PRO polypeptides may modulate glucose or free fatty acid uptake by
 CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide
 CC to factor VIIA. The PRO polypeptides can be used in assays to identify
 CC molecules involved in binding interactions. The polynucleotides encoding
 CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,
 CC transgenic or knock out animals and can be used in gene therapy.
 XX
 SO Sequence 747 AA;

Query Match 28.6%; Score 58; DB 22; Length 747;
 Best Local Similarity 27.1%; Pred. No. 4;
 Matches 13; Conservative 3; Mismatches 8; Indels 24; Gaps 1;
 QY 3 HGDRKK-----TAIGTAIAIGLGMIFTC 26
 Db 217 hgdskeslvsfamqhvrstlwtlwtgnfmslqtafaaglwltfc 264

RESULT 13
 AAB92743
 ID AAB92743 standard; Protein; 768 AA.
 XX
 AC AAB92743;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human protein sequence SEQ ID NO:11190.
 XX
 KW Human; primer: detection; diagnosis; antisense therapy; gene therapy.
 KW Homo sapiens.
 OS
 XX EPI074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-0116126.
 XX
 PR 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX
 DR WPI: 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX
 PS Claim 8; SEQ ID 11190; 2537bp + CD ROM; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX
 SO Sequence 768 AA;

Query Match 28.6%; Score 58; DB 22; Length 768;
 Best Local Similarity 27.1%; Pred. No. 4.1;
 Matches 13; Conservative 3; Mismatches 8; Indels 24; Gaps 1;
 QY 3 HGDRKK-----TAIGTAIAIGLGMIFTC 26
 Db 192 hgdskeslvsfamqhvrstlwtlwtgnfmslqtafaaglwltfc 239

RESULT 14
 AAB93204
 ID AAB93204 standard; Protein; 793 AA.
 XX
 AC AAB93204;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human protein sequence SEQ ID NO:12172.
 XX
 KW Human; primer: detection; diagnosis; antisense therapy; gene therapy.
 KW Homo sapiens.
 OS
 XX EPI074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-0116126.
 XX
 PR 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 DR WPI: 2001-318749/34.

PR Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -

PS Claim 8; SEQ ID 12172; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification, where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.

XX Sequence 793 AA;

Query Match 28.6%; Score 58; DB 22; Length 793;
 Best Local Similarity 27.1%; Pred. No. 4.3;
 Matches 13; Conservative 3; Mismatches 8; Indels 24; Gaps 1;

OY 3 HGDRKK-----TAIGLAIAIGGWIPTLC 26
 |||||
 Db 217 hgdskeslvsfamghvrstvtelwtgnfvnsiqtafaagigwiltfc 264

RESULT 15
 AAB95309

ID AAB95309 standard; Protein; 793 AA.
 AC AAB95309;

DT 26-JUN-2001 (first entry)

DE Human protein sequence SEQ ID NO:17552.

KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.

OS Homo sapiens.

PN EP1074617-A2.

PD 07-FEB-2001.

PF 28-JUL-2000; 2000EP-0116126.

PR 29-JUL-1999; 94UP-0248036.

PR 27-AUG-1999; 94JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

PA (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 DR WPI: 2001-318749/34.

PR Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -

PS Claim 8; SEQ ID 17552; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification, where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.

XX Sequence 793 AA;

Query Match 28.6%; Score 58; DB 22; Length 793;
 Best Local Similarity 27.1%; Pred. No. 4.3;
 Matches 13; Conservative 3; Mismatches 8; Indels 24; Gaps 1;

OY 3 HGDRKK-----TAIGLAIAIGGWIPTLC 26
 |||||
 Db 217 hgdskeslvsfamghvrstvtelwtgnfvnsiqtafaagigwiltfc 264

Search completed: January 10, 2002, 02:01:16
 Job time: 289 sec

Db 151 LMENRNOMIALITILGLYFTLLQASEYFE 183

RESULT 2

PCT-US91-09422-19

Sequence 19, Application PC/TUS9109422

GENERAL INFORMATION:

APPLICANT: MULVIHILL, Eileen R.
APPLICANT: Hagen, Frederick S.
APPLICANT: Houamed, Khaled M.
APPLICANT: Almers, Wolfhard
TITLE OF INVENTION: G PROTEIN-COUPLED GLUTAMATE RECEPTORS
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend
STREET: One Market Plaza, Stewart Street Tower
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105-1492

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/09422
FILING DATE: 19911212
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/672,007
FILING DATE: 18-MAR-1991
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 07/648,481
FILING DATE: 30-JAN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/626,806
FILING DATE: 12-DEC-1990
ATTORNEY/AGENT INFORMATION:
NAME: Parmelee, Steven W.
REGISTRATION NUMBER: 31,990
REFERENCE/DOCKET NUMBER: 13952-6PC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 467-9600
TELEFAX: (415) 543-5043
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 912 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US91-09422-19

Query Match 24.6%; Score 50; DB 5; Length 912;
Best Local Similarity 39.3%; Pred. No. 35;
Matches 11; Conservative 3; Mismatches 10; Indels 4; Gaps 1;

QY 4 GDRKKTATGAIATGIGTFTLCAVEY 31

Db 527 GERKTVKGMAC---WHCEPCTGYOY 550

RESULT 3

US-08-460-269C-2

Sequence 2, Application US/08460269C

GENERAL INFORMATION:

APPLICANT: CLARE, JEFFREY J.
APPLICANT: ROMANOS, MICHAEL A.
TITLE OF INVENTION: EXPRESSION OF HETEROLOGOUS PROTEIN IN
YEAST
NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

ADDRESSEE: Millen, White, Zelano & Branigan, P.C.
STREET: 2200 Clarendon Blvd., Suite 1400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22201

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,269C
FILING DATE: 02-Jun-1995
ATTORNEY/AGENT INFORMATION:
NAME: Lebovitz, Richard M.
REGISTRATION NUMBER: 37,067
REFERENCE/DOCKET NUMBER: Popov-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 243-6333
TELEFAX: (703) 243-6410
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 910 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-08-460-269C-2

Query Match 24.1%; Score 49; DB 4; Length 910;
Best Local Similarity 36.7%; Pred. No. 49;
Matches 11; Conservative 7; Mismatches 10; Indels 2; Gaps 1;

QY 2 VHGDRKKTATGAIATGIGTFTLCAVEY 31

Db 863 LRGTRAEIGLGMALRG--HSLYASYEY 890

RESULT 4

US-08-460-269C-4

Sequence 4, Application US/08460269C

GENERAL INFORMATION:

APPLICANT: CLARE, JEFFREY J.
APPLICANT: ROMANOS, MICHAEL A.
TITLE OF INVENTION: EXPRESSION OF HETEROLOGOUS PROTEIN IN
YEAST
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Millen, White, Zelano & Branigan, P.C.
STREET: 2200 Clarendon Blvd., Suite 1400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22201

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,269C
FILING DATE: 02-Jun-1995
ATTORNEY/AGENT INFORMATION:
NAME: Lebovitz, Richard M.
REGISTRATION NUMBER: 37,067
REFERENCE/DOCKET NUMBER: Popov-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 243-6333
TELEFAX: (703) 243-6410

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 911 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-08-460-269C-4

Query Match 24.1%; Score 49; DB 4; Length 911;
Best Local Similarity 36.7%; Pred. No. 49;
Matches 11; Conservative 7; Mismatches 10; Indels 2; Gaps 1;

OY 2 VHGDRKTAIGLAIGLWIFLCOAYEY 31
DB 864 LRGTRAEGLGMAALGRG--HSLVASYEY 891

RESULT 5
US-08-460-269C-6
Sequence 6, Application US/08460269C
Patent No. 6197548
GENERAL INFORMATION:
APPLICANT: CLARE, JEFFREY J.
ROMANOS, MICHAEL A.
TITLE OF INVENTION: EXPRESSION OF HETEROLOGOUS PROTEIN IN YEAST
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS: MILLER, WHITE, ZELANO & BRANNIGAN, P.C.
ADDRESSEE: MILLER, WHITE, ZELANO & BRANNIGAN, P.C.
STREET: 2200 CLARENDON BLVD., SUITE 1400
CITY: ARLINGTON
STATE: VA
COUNTRY: USA
ZIP: 22201
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,269C
FILING DATE: 02-Jun-1995
ATTORNEY/AGENT INFORMATION:
NAME: Lebovitz, Richard M.
REGISTRATION NUMBER: 37,067
REFERENCE/DOCKET NUMBER: Popov-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 243-6333
TELEFAX: (703) 243-6410
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 922 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-08-460-269C-6

Query Match 24.1%; Score 49; DB 4; Length 922;
Best Local Similarity 36.7%; Pred. No. 50;
Matches 11; Conservative 7; Mismatches 10; Indels 2; Gaps 1;

OY 2 VHGDRKTAIGLAIGLWIFLCOAYEY 31
DB 875 LRGTRAEGLGMAALGRG--HSLVASYEY 902

RESULT 6
US-08-118-441-29
Sequence 29, Application US/08118441

Patent No. 5578493
GENERAL INFORMATION:
APPLICANT: Gilliam, T. Conrad
TANZI, RUDOLPH E.
TITLE OF INVENTION: ISOLATION AND USES OF A WILSON'S DISEASE GENE
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/118,441
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/44011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525
TELEX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 1110 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-118-441-29

Query Match 24.1%; Score 49; DB 1; Length 1110;
Best Local Similarity 34.8%; Pred. No. 62;
Matches 8; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

OY 1 IVHGDRKTAIGLAIGLWIF 23
DB 863 LITGDRKTAIRALATQVGVKVF 885

RESULT 7
US-08-338-579A-29
Sequence 29, Application US/08338579A
Patent No. 6068975
GENERAL INFORMATION:
APPLICANT: Gilliam, T. Conrad
TANZI, RUDOLPH E.
TITLE OF INVENTION: ISOLATION AND USES OF A WILSON'S DISEASE GENE
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/338,579A

NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESSES:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,000
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/390,301
FILING DATE: 25-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: ADLER, REID G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: 2803-0006.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 397 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-476-000-63

Query Match 22.9%; Score 46.5; DB 1; Length 397;
Best Local Similarity 38.2%; Pred. No. 42;
Matches 13; Conservative 5; Mismatches 11; Indels 5; Gaps 2;

OY 6 RKKT-AIGLAIAGLWIFLCAVEYEIVHT 37
||| |||::|| 1: || 1: ||
DB 185 RKKAIAGISLAI--WLLTLVITPLVVKOT 215

RESULT 11
US-08-472-840-63
Sequence 63, Application US/08472840
Patent No. 5763575
GENERAL INFORMATION:
APPLICANT: SUNDELIN, JOHAN
APPLICANT: SCARBOROUGH, ROBERT M.
TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESSES:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,840
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/390,301
FILING DATE: 25-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: ADLER, REID G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: 2803-0006.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 397 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-472-840-63

Query Match 22.9%; Score 46.5; DB 1; Length 397;
Best Local Similarity 38.2%; Pred. No. 42;
Matches 13; Conservative 5; Mismatches 11; Indels 5; Gaps 2;

OY 6 RKKT-AIGLAIAGLWIFLCAVEYEIVHT 37
||| |||::|| 1: || 1: ||
DB 185 RKKAIAGISLAI--WLLTLVITPLVVKOT 215

RESULT 12
US-08-476-976-63
Sequence 63, Application US/08476976
Patent No. 5874400
GENERAL INFORMATION:
APPLICANT: SUNDELIN, JOHAN
APPLICANT: SCARBOROUGH, ROBERT M.
TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
NUMBER OF SEQUENCES: 63
CORRESPONDENCE ADDRESSES:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/476,976
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/390,301
FILING DATE: 25-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: ADLER, REID G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: 2803-0006.20
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 397 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-476-976-63

Query Match 22.9%; Score 46.5; DB 2; Length 397;
 Best Local Similarity 38.2%; Pred. No. 42;
 Matches 13; Conservative 5; Mismatches 11; Indels 5; Gaps 2;

QY 6 RKKT--AIGLAIAIGMIFTLCOAVEYEYIYHT 37
 ||| |||::|| |::| |::|
 Db 185 RKKANIAIGISLAI---WLLTLVITPLVVKQT 215

RESULT 13
 US-08-474-410-63
 ; Sequence 63, Application US/08474410
 ; Patent No. 6043212
 ; GENERAL INFORMATION:
 ; APPLICANT: SUNDELLIN, JOHAN
 ; TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS
 ; TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR
 ; NUMBER OF SEQUENCES: 63
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORRISON & FOSTER
 ; STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20006-1812
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/474,410
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/390,301
 ; FILING DATE: 25-JAN-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: ADLER, REID G.
 ; REGISTRATION NUMBER: 30,988
 ; REFERENCE/DOCKET NUMBER: 2803-0006.20
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 887-1500
 ; TELEFAX: (202) 887-0763
 ; TELEX: 90-4030
 ; INFORMATION FOR SEQ ID NO: 63:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 397 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-474-410-63

Query Match 22.9%; Score 46.5; DB 3; Length 397;
 Best Local Similarity 38.2%; Pred. No. 42;
 Matches 13; Conservative 5; Mismatches 11; Indels 5; Gaps 2;

QY 6 RKKT--AIGLAIAIGMIFTLCOAVEYEYIYHT 37
 ||| |||::|| |::| |::|
 Db 185 RKKANIAIGISLAI---WLLTLVITPLVVKQT 215

RESULT 14
 US-08-486-673B-63
 ; Sequence 63, Application US/08486673B
 ; Patent No. 6297026
 ; GENERAL INFORMATION:
 ; APPLICANT: Sundellin, Johan
 ; APPLICANT: Scarborough, Robert M.
 ; TITLE OF INVENTION: Nucleic Acids Encoding the C140 Receptor

FILE REFERENCE: 44481-5006-08-US
 ; CURRENT APPLICATION NUMBER: US/08/486,673B
 ; CURRENT FILING DATE: 1995-06-07
 ; PRIOR APPLICATION NUMBER: US 08/097,938
 ; PRIOR FILING DATE: 1993-07-26
 ; PRIOR APPLICATION NUMBER: PCT/US94/08536
 ; PRIOR FILING DATE: 1994-07-26
 ; NUMBER OF SEQ ID NOS: 63
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 63
 ; LENGTH: 397
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-08-486-673B-63

Query Match 22.9%; Score 46.5; DB 4; Length 397;
 Best Local Similarity 38.2%; Pred. No. 42;
 Matches 13; Conservative 5; Mismatches 11; Indels 5; Gaps 2;

QY 6 RKKT--AIGLAIAIGMIFTLCOAVEYEYIYHT 37
 ||| |||::|| |::| |::|
 Db 185 RKKANIAIGISLAI---WLLTLVITPLVVKQT 215

RESULT 15
 US-09-095-758-10
 ; Sequence 10, Application US/09095758
 ; Patent No. 6184016
 ; GENERAL INFORMATION:
 ; APPLICANT: PHILIPS, Mark R.
 ; TITLE OF INVENTION: HUMAN PRENYLCYSTEINE CARBOXYL
 ; TITLE OF INVENTION: METHYLTRANSFERASE, DNA ENCODING SAME, AND A METHOD OF
 ; NUMBER OF SEQUENCES: 12
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: BROWDY AND NEWMARK, P.L.L.C.
 ; STREET: 419 Seventh Street N.W., Ste. 300
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20004
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/095,758
 ; FILING DATE: 11-JUN-1998
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/049,304
 ; FILING DATE: 11-JUN-1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: YON, Allen C.
 ; REGISTRATION NUMBER: 37,971
 ; REFERENCE/DOCKET NUMBER: PHILIPS-1A
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 628-5197
 ; TELEFAX: (202) 737-3528
 ; INFORMATION FOR SEQ ID NO: 10:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 295 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-09-095-758-10

Query Match 22.7%; Score 46; DB 4; Length 295;
 Best Local Similarity 40.7%; Pred. No. 35;
 Matches 11; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

Thu Jan 10 08:15:45 2002

us-09-712-768-6.rai

Page 7

QY 8 KTAIGLAIGLWIFTLCOAYEYEI 34
| : | : | | | : | | |
Db 140 KHSGYWLAAISIGWIEFLIENFYPEI 166

Search completed: January 10, 2002, 02:01:56
Job time: 278 sec

1
1
2

PCT-US00-14973-157

us-09-712-768-7.rapm

; PRIOR APPLICATION NUMBER:US 09/540,217
 ; PRIOR FILING DATE: 2000-03-31
 ; PRIOR APPLICATION NUMBER: US 09/649,167
 ; PRIOR FILING DATE: 2000-08-23
 ; PRIOR APPLICATION NUMBER: PCT/US01/08631

1 PRIOR APPLICATION NUMBER: US 09/540,217
 2 PRIOR FILING DATE: 2000-03-31
 3 PRIOR APPLICATION NUMBER: US 09/649,167
 4 PRIOR FILING DATE: 2000-08-23
 5 PRIOR APPLICATION NUMBER: PCT/US01/086313
 6 PRIOR FILING DATE: 2001-03-30
 7 PRIOR APPLICATION NUMBER: US 09/552,929
 8 PRIOR FILING DATE: 2000-04-18
 9 PRIOR APPLICATION NUMBER: US 09/770,160
 10 PRIOR FILING DATE: 2001-01-26

```

1 PRIOR APPLICATION NUMBER: US 09/540,217
2 PRIOR FILING DATE: 2000-03-31
3 PRIOR APPLICATION NUMBER: US 09/649,167
4 PRIOR FILING DATE: 2000-08-23
5 PRIOR APPLICATION NUMBER: PCT/US01/086330
6 PRIOR FILING DATE: 2001-03-30
7 PRIOR APPLICATION NUMBER: US 09/552,929
8 PRIOR FILING DATE: 2000-04-18
9 PRIOR APPLICATION NUMBER: US 09/770,160
10 PRIOR FILING DATE: 2001-01-26
11 PRIOR APPLICATION NUMBER: PCT/US01/086566
12 PRIOR FILING DATE: 2001-04-18
13 PRIOR APPLICATION NUMBER: US 09/577,408
14 PRIOR FILING DATE: 2000-05-18
15 PRIOR APPLICATION NUMBER: PCT/US01/148221
16 PRIOR FILING DATE: 2001-05-16
17 NUMBER OF SEQ ID NOS: 3334
18 SOFTWARE: pc_files_genes version 6.0

```

```

1 PRIOR APPLICATION NUMBER: US 09/540,217
2 PRIOR FILING DATE: 2000-03-31
3 PRIOR APPLICATION NUMBER: US 09/649,167
4 PRIOR FILING DATE: 2000-08-23
5 PRIOR APPLICATION NUMBER: PCT/US01/08631
6 PRIOR FILING DATE: 2001-03-30
7 PRIOR APPLICATION NUMBER: US 09/552,929
8 PRIOR FILING DATE: 2000-04-18
9 PRIOR APPLICATION NUMBER: US 09/770,160
10 PRIOR FILING DATE: 2001-01-26
11 PRIOR APPLICATION NUMBER: PCT/US01/08656
12 PRIOR FILING DATE: 2001-04-18
13 PRIOR APPLICATION NUMBER: US 09/577,408
14 PRIOR FILING DATE: 2000-05-18
15 PRIOR APPLICATION NUMBER: PCT/US01/14822
16 PRIOR FILING DATE: 2001-05-16
17 NUMBER OF SEQ ID NOS: 3334
18 SOFTWARE: pt_FL_genes Version 6.0
19 SEQ ID NO 1004
20 LENGTH: 153

```

```

: - ORANISM: Homo sapiens
US-60-324-631-1004

alignment_scores:
    Quality: 65.00
    Ratio: 4.062
    Percent Similarity: 55.172
    Percent Identity: 55.172

alignment_block:
US-09-712-768-7/rev x US-60-324-631-1004

Align seg 1/1 to: US-60-324-631-1004 from
73 AACCGACGCTGTTTGGCCGACATCGCACCGC
|||||
124 AsnAtgAlgaValValLeuProArgSerHisArg
40 .....GCAGGATGCGGATGACGCACAGCAGCA
|||||
140 cAlaGlnValAGlyCysLeuProGlyAlaGlnAla

seq_name: /cgn2_6/prodata/2/paa/US092_COMB

seq_documentation_block:
; Sequence 18747, Application US/09352991A
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMPLIFICATION
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,995
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18747
; LENGTH: 230
; TYPE: PRT
; ORANISM: Pseudomonas aeruginosa
US-09-252-991A-18747

alignment_scores:
    Quality: 64.50
    Ratio: 3.395
    Percent Similarity: 70.370
    Percent Identity: 70.370

alignment_block:
US-09-712-768-7/rev x US-09-252-991A-18747

```

SS EUI

Align seg 1/1 to: US-09-252-991A-18747 from: 1 to: 230

```
82 CCATCTCGA.....AACGACGTGCTGTTTGGGACATC 48
|||||
52 ProSerArgAlaHisThrGlySerSerArgGlySerAlaProThrSe 68
||:|||||
47 GCACCGCGCAGATCGGATCAGGACACA 17
||:|||||
68 ArgArgAlaAlaCysGlySerMetArgPro 78
```

seq_name: /cgn2_6/prodata/2/paa/PCTUS_COMB.pep:PCT-US00-06013-142

seq_documentation_block:

```
; Sequence 142, Application PC/TUS0006013
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: 50 Human Secreted Proteins
; FILE REFERENCE: PS508PCT
; CURRENT APPLICATION NUMBER: PCT/US00/06013
; CURRENT FILING DATE: 2000-03-09
; EARLIER APPLICATION NUMBER: 60/125,360
; EARLIER FILING DATE: 1999-03-19
; EARLIER APPLICATION NUMBER: 60/138,626
; EARLIER FILING DATE: 1999-06-11
; EARLIER APPLICATION NUMBER: 60/168,662
; EARLIER FILING DATE: 1999-12-03
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 142
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (41)
; OTHER INFORMATION: Xaa equals stop translation
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (50)
; OTHER INFORMATION: Xaa equals stop translation
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (59)
; OTHER INFORMATION: Xaa equals stop translation
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (60)
; OTHER INFORMATION: Xaa equals stop translation
PCT-US00-06013-142
```

alignment_scores:

Quality:	64.00	Length:	28
Ratio:	3.200	Gaps:	0
Percent Similarity:	71.429	Percent Identity:	50.000

alignment_block:

US-09-712-768-7 x PCT-US00-06013-142 ..

Align seg 1/1 to: PCT-US00-06013-142 from: 1 to: 62

```
4 TCGATCTTCCTGCTGCTGCTGATCCGATCCTGCGGCGTGCGATGC 53
||| ||||| ::|||::||| ||| :::::
13 SerThrPheLeuThrIleCysPheIleArgGlnLeuIlePheHisPheTh 29
|||||
54 GCACAAACAGCACGCTCGGTTTCGAGATGCCGCA 87
:::||||| ||||| |||||
29 rSerLyHisHisPheGlyPheGlnAlaAlaAla 40
```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2002, 01:51:17 ; Search time 18.96 Seconds
(without alignments)
152.671 Million cell updates/sec

Title: US-09-712-768-6
Perfect score: 203
Sequence: 1 IYHGDRKRTAIGLAIAIGLGMIFTLCAQAYEYEVHT 38

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: PIR1:
2: PIR2:
3: PIR3:
4: PIR4:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	87	42.9	260	1	OTXL3	cytochrome-c oxida
2	84	41.4	261	2	T10978	cytochrome-c oxida
3	83	40.9	261	2	T11863	cytochrome-c oxida
4	83	40.9	261	2	T11369	cytochrome-c oxida
5	82.5	40.6	266	2	C45164	cytochrome-c oxida
6	81	39.9	261	2	T11283	cytochrome-c oxida
7	81	39.9	274	2	S03807	cytochrome-c oxida
8	80	39.4	261	2	B50396	cytochrome-c oxida
9	80	39.4	261	2	H34012	cytochrome-c oxida
10	80	39.4	261	2	G34012	cytochrome-c oxida
11	80	39.4	261	2	A30396	cytochrome-c oxida
12	80	39.4	261	2	D30396	cytochrome-c oxida
13	80	39.4	261	2	C30396	cytochrome-c oxida
14	80	39.4	261	2	T09953	cytochrome-c oxida
15	79	38.9	261	1	OTB03	cytochrome-c oxida
16	79	38.9	261	1	OTCA3	cytochrome-c oxida
17	79	38.9	261	2	S47876	cytochrome-c oxida
18	79	38.9	261	2	T11486	cytochrome-c oxida
19	77	37.9	261	2	G58850	cytochrome-c oxida
20	77	37.9	261	2	G58850	cytochrome-c oxida
21	77	37.9	261	2	S41841	cytochrome-c oxida
22	76	37.4	261	2	S26157	cytochrome-c oxida
23	76	37.4	261	2	T11499	cytochrome-c oxida
24	76	37.4	261	2	T11253	cytochrome-c oxida
25	76	37.4	261	2	T11408	cytochrome-c oxida
26	75	36.9	261	2	T11081	cytochrome-c oxida
27	75	36.9	261	2	S68134	cytochrome-c oxida
28	74	36.5	261	2	T11512	cytochrome-c oxida
29	74	36.5	261	2	T11512	cytochrome-c oxida

30	74	36.5	261	2	T11447	cytochrome-c oxida
31	74	36.5	261	2	T11540	cytochrome-c oxida
32	74	36.5	261	2	T11146	cytochrome-c oxida
33	74	36.5	261	2	T11434	cytochrome-c oxida
34	74	36.5	261	2	T45556	cytochrome-c oxida
35	72	35.5	261	2	S60277	cytochrome-c oxida
36	72	35.5	261	2	T11826	cytochrome-c oxida
37	72	35.5	261	2	T11173	cytochrome-c oxida
38	71	35.0	261	2	T11306	cytochrome-c oxida
39	71	35.0	262	2	T11883	cytochrome-c oxida
40	71	35.0	262	2	S62760	cytochrome-c oxida
41	71	35.0	270	2	S10193	cytochrome-c oxida
42	70	34.5	261	2	T11107	cytochrome-c oxida
43	70	34.5	262	2	T11390	cytochrome-c oxida
44	70	34.5	262	2	H71390	cytochrome-c oxida
45	69	34.0	260	2	T11798	cytochrome-c oxida

ALIGNMENTS

RESULT 1

OTXL3
cytochrome-c oxidase (EC 1.9.3.1) chain III - African clawed frog mitochondrion
N:Alternate names: cytochrome a3 polypeptide III; cytochrome a3 polypeptide III
C:Species: mitochondrion Xenopus laevis (African clawed frog)
C:Date: 28-Feb-1986 #sequence_Revision 28-Feb-1986 #text_change 07-Dec-1999
C:Accession: A00486
R:Roel, B.A.; Ma, D.P.; Wilson, R.K.; Wong, J.F.H.
J: Biol. Chem. 260, 9759-9774, 1985
A:Title: The complete nucleotide sequence of the Xenopus laevis mitochondrial genome.
A:Reference number: A00155; M01D:85261388
A:Accession: A00486
A:Molecule type: DNA
A:Residues: 1-260 <ROE>
A:Cross-references: GB:M10217; GB:X01600; GB:X01601; GB:X02890; NID:9343717; PIDN:AAA
A:Genetic code: SGC1
A:Gene: coll1
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGC1
C:Superfamily: cytochrome-c oxidase chain III
C:Keywords: electron transfer; membrane-associated complex; mitochondrial inner membr
F:16-34/Domain: transmembrane helix #status predicted <TR01>
F:41-66/Domain: transmembrane helix #status predicted <TR02>
F:73-105/Domain: transmembrane helix #status predicted <TR03>
F:129-152/Domain: transmembrane helix #status predicted <TR04>
F:156-183/Domain: transmembrane helix #status predicted <TR05>
F:191-223/Domain: transmembrane helix #status predicted <TR06>
F:233-255/Domain: transmembrane helix #status predicted <TR07>

Query Match 42.9%; Score 87; DB 1; Length 260;
Best Local Similarity 51.4%; Pred. No. 0.00022;
Matches 19; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

QY 1 IYHGDRKRTAIGLAIAIGLGMIFTLCAQAYEYEVHT 37
DB 151 IMHGDREKRTAIGLAIAIGLGMIFTLCAQAYEYEVHT 187

RESULT 2
T10978
cytochrome-c oxidase (EC 1.9.3.1) chain III - pig mitochondrion
C:Species: mitochondrion Sus scrofa domestica (domestic pig)
C:Date: 16-Jul-1999 #sequence_Revision 16-Jul-1999 #text_change 20-Jun-2000
C:Accession: T10978; T11876; C45954
R:Lin, C.S.; Liu, C.Y.; Sun, Y.L.; Chang, L.C.; Cheng, I.C.; Yang, P.C.; Mao, S.J.T.;
submitted to the EMBL Data Library, November 1997
A:Description: Complete nucleotide sequence of the porcine mitochondrial genome.
A:Reference number: Z17237
A:Accession: T10978
A:Status: preliminary; translated from GB/EMBL/DDAJ
A:Molecule type: DNA

Query Match	40.98;	Score 83;	DB 2;	Length 261;
Best Local Similarity	51.48;			
Matches 19;	Conservative	3.	Microbial	

cytochrome-c oxidase (EC 1.9.3.1) chain III - Rhodobacter sphaeroides
C:Species: Rhodobacter sphaeroides
C:Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 17-Mar-1999
C:Accession: C45164
R:Cao, J.; Hostler, J.; Shapleigh, J.; Revzin, A.; Ferguson-Miller, S.
J Biol. Chem. 267, 24273-24278, 1992
Title: Cytochrome aa-3 of Rhodobacter sphaeroides

```

A:Reference number: A4S164; MUID:93077508
A:Accession: C4S164
A>Status: Preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-266 <C>AO>
C:Genetics:
A:Gene: coliI
C:Superfamily: cytochrome-c oxidase chain III
C:Keywords: electron transfer; membrane-associated complex; oxidoreductase; transmembr
Query Match          40.6%; Score 82.5; DB 2; Length 266;
Best Local Similarity 56.2%; Pred No. 0.00094;
Matches 18; Conservative 6; Mismatches 7; Indels 1; Gaps 1;
QY      1 IVH-GDKRTTAIGLAIATGIGTWTTCATVEY 31
       :|||:::| |||:::||:::||::|||
Db      155 LVHENNRDVAWGALATATAGLTFTVFAEY 186

```

RESULT 6
T11293
cytochrome-c oxidase (EC 1.9.3.1) chain III - Raja radiata mitochondrion
C:Species: mitochondrion Raja radiata
C:date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 11-May-2000
R:Rasmussen, A.S.; Arnason, U.
Proc. Natl. Acad. Sci. U.S.A. 96, 2177-2182, 1999
A:title: Molecular studies suggest that cartilaginous fishes have an apical position
A:reference_number: 217259; MUID:99162577
A:accession: T11293
A:status: preliminary; translated from GB/EMBL/DBJ
A:molecule_type: DNA
A:residues: 1-261 <RAS>
A:cross-references: EMBL:AF106038; NID:g4406269; PID:g4406276; PIDN:AAD19335.1
A:genetics:
A:gene: COX-III
A:genome: mitochondrion

A:Genetic code: SGC1
C:Superfamily: cytochrome-c oxidase chain III
C:Keywords: electron transfer; membrane-associated complex; mitochondrial inner membrane

Query Match 39.4%; Score 80; DB 2; Length 261;
Best Local Similarity 45.9%; Pred. No. 0.002;
Matches 17; Conservative 5; Mismatches 15; Indels 0; Gaps 0;

QY 1 IVHGDRKKTALGAIATGIMIFLCOAYEYEVHT 37
| : ||| : | : ||| : | : ||| : |
Db 151 IMEGRKQTIALTLTLTGFTFLQGMETYEAPFT 187

RESULT 11
A30396
cytochrome-c oxidase (EC 1.9.3.1) chain III - sockeye salmon mitochondrion
C:Species: mitochondrion Oncorhynchus nerka (sockeye salmon)
C>Date: 22-Jun-1990 #sequence_revision 22-Jun-1990 #text_change 07-Dec-1999
C:Accession: A30396
R:Thomas, W.K.; Beckenbach, A.T.
J. Mol. Evol. 29, 233-245, 1989
A:Title: Variation in salmonid mitochondrial DNA: evolutionary constraints and mechanism
A:Reference number: A30396; MUID:89382658
A:Accession: A30396
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-261 <THO>
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGC1
C:Superfamily: cytochrome-c oxidase chain III
C:Keywords: electron transfer; membrane-associated complex; mitochondrial inner membrane

Query Match 39.4%; Score 80; DB 2; Length 261;
Best Local Similarity 45.9%; Pred. No. 0.002;
Matches 17; Conservative 5; Mismatches 15; Indels 0; Gaps 0;

QY 1 IVHGDRKKTALGAIATGIMIFLCOAYEYEVHT 37
| : ||| : | : ||| : | : ||| : |
Db 151 IMEGRKQTIALTLTLTGFTFLQGMETYEAPFT 187

RESULT 12
D30396
cytochrome-c oxidase (EC 1.9.3.1) chain III - cutthroat trout mitochondrion
C:Species: mitochondrion Salmo clarki (cutthroat trout)
C>Date: 22-Jun-1990 #sequence_revision 22-Jun-1990 #text_change 07-Dec-1999
C:Accession: D30396
R:Thomas, W.K.; Beckenbach, A.T.
J. Mol. Evol. 29, 233-245, 1989
A:Title: Variation in salmonid mitochondrial DNA: evolutionary constraints and mechanism
A:Reference number: A30396; MUID:89382658
A:Accession: D30396
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-261 <THO>
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGC1
C:Superfamily: cytochrome-c oxidase chain III
C:Keywords: electron transfer; membrane-associated complex; mitochondrial inner membrane

Query Match 39.4%; Score 80; DB 2; Length 261;
Best Local Similarity 45.9%; Pred. No. 0.002;
Matches 17; Conservative 5; Mismatches 15; Indels 0; Gaps 0;

QY 1 IVHGDRKKTALGAIATGIMIFLCOAYEYEVHT 37
| : ||| : | : ||| : | : ||| : |
Db 151 IMEGRKQTIALTLTLTGFTFLQGMETYEAPFT 187

RESULT 13
C30396
cytochrome-c oxidase (EC 1.9.3.1) chain III - chinook salmon mitochondrion
C:Species: mitochondrion Oncorhynchus tshawytscha (chinook salmon)
C>Date: 22-Jun-1990 #sequence_revision 22-Jun-1990 #text_change 07-Dec-1999
C:Accession: C30396
R:Thomas, W.K.; Beckenbach, A.T.
J. Mol. Evol. 29, 233-245, 1989
A:Title: Variation in salmonid mitochondrial DNA: evolutionary constraints and mechanism
A:Reference number: A30396; MUID:89382658
A:Accession: C30396
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-261 <THO>
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGC1
C:Superfamily: cytochrome-c oxidase chain III
C:Keywords: electron transfer; membrane-associated complex; mitochondrial inner membrane

Query Match 39.4%; Score 80; DB 2; Length 261;
Best Local Similarity 45.9%; Pred. No. 0.002;
Matches 17; Conservative 5; Mismatches 15; Indels 0; Gaps 0;

QY 1 IVHGDRKKTALGAIATGIMIFLCOAYEYEVHT 37
| : ||| : | : ||| : | : ||| : |
Db 151 IMEGRKQTIALTLTLTGFTFLQGMETYEAPFT 187

RESULT 14
T09953
cytochrome-c oxidase (EC 1.9.3.1) chain III - Atlantic salmon mitochondrion
C:Species: mitochondrion Salmo salar (Atlantic salmon)
C>Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 07-Dec-1999
C:Accession: T09953; B61025
R:Hurst, C.D.; Bartlett, S.E.; Bruce, I.J.; Davidson, W.S.
submitted to the EMBL Data Library, October 1998
A:Description: The complete nucleotide sequence of the mitochondrial DNA of the Atlantic salmon.
A:Reference number: 216904
A:Accession: T09953
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-261 <HUB>
A:Cross-references: EMBL:U12143; EMBL:U12144; EMBL:U12145; EMBL:U12146; NID:g3775976;
A:Experimental source: liver
R:Davidson, W.S.; Birt, T.P.; Green, J.M.
Genome 32, 340-342, 1989
A:Title: Organisation of the mitochondrial genome from Atlantic salmon (Salmo salar)
A:Reference number: A61025; MUID:89306582
A:Accession: B61025
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 135-139, 'L', 141-158, 'M', 160-163, 'L', 165-182, 'G', 185-191, 'Y', 194-215, 'T'
C:Genetics:
A:Genome: mitochondrion
A:Genetic code: SGC1
C:Superfamily: cytochrome-c oxidase chain III
C:Keywords: electron transfer; membrane-associated complex; mitochondrial inner membrane

Query Match 39.4%; Score 80; DB 2; Length 261;
Best Local Similarity 45.9%; Pred. No. 0.002;
Matches 17; Conservative 5; Mismatches 15; Indels 0; Gaps 0;

QY 1 IVHGDRKKTALGAIATGIMIFLCOAYEYEVHT 37
| : ||| : | : ||| : | : ||| : |
Db 151 IMEGRKQTIALTLTLTGFTFLQGMETYEAPFT 187

RESULT 15
OTB03

cytochrome-c oxidase (EC 1.9.3.1) chain III [validated] - bovine mitochondrion
 C:Species: mitochondrion Bos primigenius taurus (cattle)
 C:Date: 18-Aug-1982 #sequence_revision 18-Aug-1982 #text_change 15-Sep-2000
 C:Accession: A00483
 R:Anderson, S.; de Bruijn, M.H.L.; Coulson, A.R.; Eperon, I.C.; Sanger, F.; Young, I.G.
 J. Mol. Biol. 156, 683-717, 1982
 A:Title: Complete sequence of bovine mitochondrial DNA. Conserved features of the mammal
 A:Reference number: A00152; MUID:83010260
 A:Accession: A00483
 A:Molecule type: DNA
 A:Residues: 1-261 <AND>
 A:Cross-references: GB:J01394; NID:9336430; EMBL:V00654; NID:912800; PID:912807
 R:Tsukihara, T.; Aoyama, H.; Yamashita, E.; Tomizaki, T.; Yamaguchi, H.; Shinzawa-Itch,
 submitted to the Brookhaven Protein Data Bank, April 1996
 A:Reference number: A67451; PDB:1OCC
 A:Contents: annotation: X-ray crystallography, 2.8 angstroms, residues 1-261
 R:Tsukihara, T.; Aoyama, H.; Yamashita, E.; Tomizaki, T.; Yamaguchi, H.; Shinzawa-Itch, R
 Science 272, 1136-1144, 1996
 A:Title: The whole structure of the 13-subunit oxidized cytochrome c oxidase at 2.8 ang
 A:Reference number: A57981; MUID:96216288
 A:Contents: annotation: X-ray crystallography, 2.8 angstroms
 C:Genetics:
 A:Gene: coxi
 A:Genome: mitochondrion
 A:Genetic code: SGC1
 C:Complex: part of a 13 chain complex spanning the inner mitochondrial membrane and cons
 Vb (see PIR:OGBO6), Via (see PIR:OGBO6), Vtb (see PIR:OGBO7), Vtc (see PIR:OGBO6C), Vt
 iners within the mitochondrial inner-membrane
 C:Function:
 A:Description: the cytochrome-c oxidase complex catalyzes the oxidation of four molecu
 ns from the mitochondrial matrix producing two molecules of water and lowering the conce
 A:pathway: oxidative phosphorylation; respiratory chain
 A>Note: chain III may help bind cytochrome c on the intracristal side of the inner-membr
 C:Superfamily: cytochrome-c oxidase chain III
 C:Keywords: electron transfer; membrane-associated complex; mitochondrial inner membrane
 F:1-15/Domain: mitochondrial matrix #status experimental <MM1>
 F:16-34/Domain: transmembrane helix #status experimental <TR01>
 F:35-40/Domain: intracristal #status experimental <ITC1>
 F:41-66/Domain: transmembrane helix #status experimental <TR02>
 F:67-72/Domain: mitochondrial matrix #status experimental <TR03>
 F:73-105/Domain: transmembrane helix #status experimental <MM2>
 F:106-128/Domain: intracristal #status experimental <ITC2>
 F:129-152/Domain: transmembrane helix #status experimental <TR04>
 F:153-155/Domain: mitochondrial matrix #status experimental <MM3>
 F:156-183/Domain: transmembrane helix #status experimental <TR05>
 F:184-190/Domain: intracristal #status experimental <ITC3>
 F:191-223/Domain: transmembrane helix #status experimental <TR06>
 F:224-232/Domain: mitochondrial matrix #status experimental <MM4>
 F:233-256/Domain: transmembrane helix #status experimental <TR07>
 F:257-261/Domain: intracristal #status experimental <ITC4>

Query Match 38.9%; Score 79; DB 1; Length 261;
 Best Local Similarity 51.4%; Pred. No. 0.0028;
 Matches 19; Conservative 2; Mismatches 16; Indels 0; Gaps 0;

OY 1 IYHSDRKTATIGLAIATIGLWFTLCQATVEYIYHT 37
 :: ||||| | | | | | | | | | | | | | | | |
 Db 151 LMEGRKMLQALFTITLGVFTLLQASEYYEAPPT 187

Search completed: January 10, 2002, 02:00:24
 Job time: 547 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2002, 01:59:48 ; Search time 16.73 Seconds
(without alignments)
83.279 Million cell updates/sec

Title: US-09-712-768-6
Perfect score: 203
Sequence: 1 IVHGDRKRTAIGLAIGMIFTLQAYEYEVHT 38

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues
Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	87	42.9	260	1	COX3_XENLA
2	84	41.4	261	1	COX3_PIG
3	83	40.9	261	1	COX3_EQUAS
4	83	40.9	261	1	COX3_HORSE
5	82	40.4	261	1	COX3_CAVAU
6	81	39.9	273	1	COX3_PARDE
7	80	39.4	261	1	COX3_ONCCL
8	80	39.4	261	1	COX3_ONCME
9	80	39.4	261	1	COX3_ONCNE
10	80	39.4	261	1	COX3_SALSA
11	79	38.9	261	1	COX3_BOVIN
12	79	38.9	261	1	COX3_CYPCA
13	79	38.9	261	1	COX3_RABIT
14	79	38.9	261	1	COX3_DIDMA
15	77	37.9	261	1	COX3_BALMU
16	77	37.9	261	1	COX3_BALPH
17	77	37.9	261	1	COX3_ORNAN
18	76	37.4	261	1	COX3_CANPA
19	76	37.4	261	1	COX3_CERSI
20	76	37.4	261	1	COX3_FELCA
21	76	37.4	261	1	COX3_HALGR
22	76	37.4	261	1	COX3_PHOVI
23	76	37.4	261	1	COX3_RHITU
24	76	37.4	261	1	COX3_TIRAM
25	74	36.5	261	1	COX3_DASNO
26	74	36.5	261	1	COX3_MACRO
27	74	36.5	261	1	COX3_PAPPA
28	74	36.5	261	1	COX3_SOUAC
29	73	36.0	261	1	COX3_HIPMA
30	72	35.5	261	1	COX3_CROMA
31	72	35.5	261	1	COX3_GADNO
32	71	35.0	261	1	COX3_AEPME
33	71	35.0	261	1	COX3_ANTLOPE

34	71	35.0	261	1	COX3_ANTWR	047701 antidorcas
35	71	35.0	261	1	COX3_CEPNA	047693 cephalophus
36	71	35.0	261	1	COX3_DAMU	047694 damalisca
37	71	35.0	261	1	COX3_GAZCU	047708 gazella cuv
38	71	35.0	261	1	COX3_GAZDA	048374 gazella dam
39	71	35.0	261	1	COX3_GAZDO	048308 gazella dor
40	71	35.0	261	1	COX3_GAZGA	048346 gazella gaz
41	71	35.0	261	1	COX3_GAZGR	047703 gazella gra
42	71	35.0	261	1	COX3_GAZLE	047709 gazella lep
43	71	35.0	261	1	COX3_GAZRU	047705 gazella lep
44	71	35.0	261	1	COX3_GAZSA	047710 gazella sau
45	71	35.0	261	1	COX3_GAZSU	048316 gazella sub

ALIGNMENTS

RESULT 1
COX3_XENLA STANDARD; PRT; 260 AA.
ID COX3_XENLA
AC P00419;
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE CYTOCHROME C OXIDASE POLYPEPTIDE III (EC 1.9.3.1).
GN COIII.
OS Xenopus laevis (African clawed frog).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85261386; PubMed=4019494;
RA Roe B.A., Ma D.-P., Wilson R.K., Wong J.F.-H.;
RT "The complete nucleotide sequence of the Xenopus laevis mitochondrial
genome.";
RL J. Biol. Chem. 260:9759-9774(1985).
CC -I- PUNCTION: SUBUNIT I, II, AND III FORM THE FUNCTIONAL CORE OF
THE ENZYME COMPLEX.
CC -I- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) -> 2 H(2)O +
4 FERRICYTOCHROME C.
CC -I- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC EMBL: M10217; AAA6464.1; .
DR PIR: A00486; OYX13.
DR HSSP: P00415; IOCC.
DR InterPro: IPR000298; CytC_oxdse_III.
DR Pfam: PF00510; COX3; 1.
DR PROSITE: PS00253; COX3; 1.
KW Oxidoreductase; Mitochondrion; Transmembrane.
SQ SEQUENCE 260 AA; 29593 MW; FA736567CEAB3826 CRC64;

Query Match 42.9%; Score 87; DB 1; Length 260;
Best Local Similarity 51.4%; Pred. No. 0.00036;
Matches 19; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

QY 1 IVHGDRKRTAIGLAIGMIFTLQAYEYEVHT 37
DB 151 IHGDRKRAIQSLTILILGLYFALQAMEYEAFT 187
RESULT 2

```

COX3_PIG
ID COX3_PIG STANDARD; PRT; 261 AA.
AC Q35916; 079879;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE CYTOCHROME C OXIDASE POLYPEPTIDE III (EC 1.9.3.1).
GN MTCO3 OR COIII.
OS Sus scrofa (Pig).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98404150; PubMed=9732457;
RA Drsing B.M., Arnason U.;
RT "The complete mitochondrial DNA sequence of the pig (Sus scrofa).";
RT J. Mol. Evol. 47:302-306(1998).
RN [2]
RP SEQUENCE OF 1-30 FROM N.A.
RX MEDLINE=86295557; PubMed=3017295;
RA Watanabe T., Hayashi Y., Kimura J., Yasuda Y., Saitou N., Tomita T.,
RA Ogawara N.;
RT "Pig mitochondrial DNA: polymorphism, restriction map orientation,
RT and sequence data.";
RL Biochem. Genet. 24:385-396(1986).
CC -1- FUNCTION: SUBUNIT I, II, AND III FORM THE FUNCTIONAL CORE OF
CC THE ENZYME COMPLEX.
CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
CC 4 FERRICYTOCHROME C.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ002189; CAA05235.1; ALT_INIT.
DR EMBL: M26139; AAA32032.1; -.
DR InterPro: IPR000298; Cytc_oxdse_III.
DR Pfam: PF00510; COX3; 1.
DR PROSITE: PS50253; COX3; 1.
KW Oxidoreductase; Mitochondrion; Transmembrane.
FT CONFLICT 22 24 LSA -> YSG (IN REF. 2).
FT SEQUENCE 261 AA; 29728 MW; 07E93FCA7FC1784B CRC64;
SQ

```

```

Query Match 41.4%; Score 84; DB 1; Length 261;
Best Local Similarity 51.4%; Pred. No. 0.00087;
Matches 19; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

```

```

OY 1 IVHGDRKRTAIGLAIGWIFTLQAYEYETVHT 37
   ::||| |::||| |::||| |::||| |
Db 151 LMEDGRKRMIGALSTITALGYFTLLQASERYEASF 187

RESULT 3
COX3_EQUUS STANDARD; PRT; 261 AA.
AC P92481;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE CYTOCHROME C OXIDASE POLYPEPTIDE III (EC 1.9.3.1).
GN MTCO3 OR COIII.
OS Equus asinus (Donkey).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.

```

```

OX NCBI_TaxID=9793;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=97032591; PubMed=8875857;
RA Xu X., Gullberg A., Arnason U.;
RT "The complete mitochondrial DNA (mtDNA) of the donkey and mtDNA
RT comparisons among four closely related mammalian species pairs.";
RL J. Mol. Evol. 43:438-463(1996).
CC -1- FUNCTION: SUBUNIT I, II, AND III FORM THE FUNCTIONAL CORE OF
CC THE ENZYME COMPLEX.
CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
CC 4 FERRICYTOCHROME C.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X97337; CAA66020.1; -.
DR InterPro: IPR000298; Cytc_oxdse_III.
DR Pfam: PF00510; COX3; 1.
DR PROSITE: PS50253; COX3; 1.
KW Oxidoreductase; Mitochondrion; Transmembrane.
FT CONFLICT 261 AA; 29761 MW; EB33C3F0BEA40769 CRC64;
SQ

```

```

Query Match 40.9%; Score 83; DB 1; Length 261;
Best Local Similarity 51.4%; Pred. No. 0.0012;
Matches 19; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

```

```

OY 1 IVHGDRKRTAIGLAIGWIFTLQAYEYETVHT 37
   ::||| |::||| |::||| |::||| |
Db 151 LMEDGRKRMIGALSTITALGYFTLLQASERYEASF 187

```

```

RESULT 4
COX3_HORSE STANDARD; PRT; 261 AA.
AC P48661;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE CYTOCHROME C OXIDASE POLYPEPTIDE III (EC 1.9.3.1).
GN MTCO3 OR COIII.
OS Equus caballus (Horse).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95047450; PubMed=7958969;
RA Xu X., Arnason U.;
RT "The complete mitochondrial DNA sequence of the horse, Equus
RT caballus: extensive heteroplasmy of the control region.";
RL Gene 148:357-362(1994).
CC -1- FUNCTION: SUBUNIT I, II, AND III FORM THE FUNCTIONAL CORE OF
CC THE ENZYME COMPLEX.
CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
CC 4 FERRICYTOCHROME C.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

or send an email to license@lsb-sib.ch).

CC or send an email to license@lsb-sib.ch).
 CC EMBL: X79547; CAA56085.1; -
 DR HSP: P00415; I0CC. CytC_oxdase_III.
 DR InterPro: IPR000298; CytC_oxdase_III.
 DR Pfam: PF00510; COX3; 1.
 DR PROSITE: PS50253; COX3; 1.
 KW Oxidoreductase; Mitochondrion; Transmembrane.
 SQ SEQUENCE 261 AA; 29779 MW; F09368EBA4BFA4CC9 CRC64;

Query Match 40.9%; Score 83; DB 1; Length 261;
 Best Local Similarity 51.4%; Pred. No. 0.0012;
 Matches 19; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

OY 1 IVHGRKTAIGLAIGLWITLQAEYVEIVHT 37
 Db 151 IMEGRRKMLGLFTTISGLVFTLLQASEYEASFT 187

RESULT 5
 COX3_CARAU STANDARD; PRT; 261 AA.

AC 096133; 078685;
 DT 13-JUL-1998 (Rel. 36, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE CYTOCHROME C OXIDASE POLYPEPTIDE III (EC 1.9.3.1).
 GN COXIII OR COIII.
 OS Carassius auratus (goldfish).
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
 OC Cypriniformes; Cyprinidae; Cyprininae; Carassius.
 OX NCBI_TaxID=7957;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AZ3; TISSUE=Oocyte;
 RA Murakami M., Yamashita Y., Fujitani H.;
 RT "The complete sequence of mitochondrial genome from a gynogenetic
 RT triploid 'gibbuna' (Carassius auratus langsdorffii).";
 RL Zool. Sci. 15:335-337(1998).
 RN [2]
 RP SEQUENCE OF 77-261 FROM N.A.
 RA Lin X.-W.;
 RL Submitted (MAY-1996) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: SUBUNIT I, II, AND III FORM THE FUNCTIONAL CORE OF
 CC THE ENZYME COMPLEX.
 CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
 CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@lsb-sib.ch).
 CC -----
 CC EMBL: AB006953; BAA31244.1; -
 DR EMBL: U57388; AAA99828.1; -
 DR InterPro: IPR000298; CytC_oxdase_III.
 DR Pfam: PF00510; COX3; 1.
 DR PROSITE: PS50253; COX3; 1.
 KW Oxidoreductase; Mitochondrion; Transmembrane.
 FT CONFLICT 171 171 F -> L (IN REF. 2).
 FT SEQUENCE 261 AA; 29674 MW; C8BDB3AD5BE184FF CRC64;

Query Match 40.4%; Score 82; DB 1; Length 261;
 Best Local Similarity 48.6%; Pred. No. 0.0016;
 Matches 18; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

OY 1 IVHGRKTAIGLAIGLWITLQAEYVEIVHT 37
 Db 151 IMEGRRKMLGLFTTISGLVFTLLQASEYEASFT 187

RESULT 6
 COX3_CARAU STANDARD; PRT; 273 AA.

AC 13-AUG-1987 (Rel. 05, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE CYTOCHROME C OXIDASE POLYPEPTIDE III (EC 1.9.3.1) (CYTOCHROME AA3
 DE SUBUNIT 3) (OXIDASE AA(3) SUBUNIT 3).
 GN CTAE OR COIII.
 OS Paracoccus denitrificans.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhodospirillum rubrum;
 OC Paracoccus.
 OX NCBI_TaxID=266;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SI1657;
 RA Raitio M., Jalli T., Saraste M.;
 RT "Isolation and analysis of the genes for cytochrome c oxidase in
 RT Paracoccus denitrificans.";
 RL EMBO J. 6:2825-2833(1987).
 RN [2]
 RP SEQUENCE OF 1-143 FROM N.A.
 RC STRAIN=SI1657;
 RX MEDLINE=87005242; PubMed=3019767;
 RA Saraste M., Raitio M., Jalli T., Peramäki A.;
 RT "A gene in Paracoccus for subunit III of cytochrome oxidase.";
 RL FEBS Lett. 206:154-156(1986).
 RN [3]
 RP SEQUENCE OF 1-15.
 RX MEDLINE=8616732; PubMed=2832167;
 RA Haltia T., Pustinen A., Finel M.;
 RT "The Paracoccus denitrificans cytochrome aa3 has a third subunit.";
 RL Eur. J. Biochem. 172:543-546(1988).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
 RX MEDLINE=95379947; PubMed=7651515;
 RA Iwata S., Ostermeier C., Ludwig B., Michel H.;
 RT "Structure at 2.8-A resolution of cytochrome c oxidase from
 RT Paracoccus denitrificans.";
 RL Nature 376:660-669(1995).
 CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
 CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN; INNER MEMBRANE.
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@lsb-sib.ch).
 CC -----
 CC EMBL: X04406; CAA27995.1; -
 DR EMBL: X05828; CAA29272.1; -
 DR PIR: S03807; S03807.
 DR PIR: A24371; A24371.
 DR HSP: P00415; I0CC.
 DR InterPro: IPR000298; CytC_oxdase_III.
 DR Pfam: PF00510; COX3; 1.
 KW Oxidoreductase; Transmembrane; Inner membrane.
 FT INIT_MET 0
 FT DOMAIN 1 14 CYTOPLASMIC.
 FT TRANSMEM 15 35 PERIPLASMIC.
 FT DOMAIN 36 47

```

FT TRANSMEM 48 76 CYTOPLASMTIC.
FT DOMAIN 77 78
FT TRANSMEM 79 114
FT DOMAIN 115 138 PERIPLASMTIC.
FT TRANSMEM 139 165
FT DOMAIN 166 167 CYTOPLASMTIC.
FT TRANSMEM 168 196
FT DOMAIN 197 202 PERIPLASMTIC.
FT TRANSMEM 203 236
FT DOMAIN 237 243 CYTOPLASMTIC.
FT TRANSMEM 244 273
SQ SEQUENCE 273 AA; 30655 MW; 0831BD966AE3C7D6 CRC64;

```

```

Query Match
Best Local Similarity 39.4%; Score 81; DB 1; Length 273;
Matches 17; Conservative 4; Mismatches 10; Indels 0; Gaps 0;
QY 1 IVHDKRRTAIGLAIGMIFTLQAYEYEVHT 31
Db 163 VLEGRKTTINGLIVAVILGVCFTGLQAYEY 193

```

```

RESULT 7
ID COX3_ONCCL STANDARD; PRT; 261 AA.
AC P20684;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DE CYTOCHROME C OXIDASE POLYPEPTIDE III (EC 1.9.3.1).
GN COXIII OR COIII.
OS Oncorhynchus clarki (Cutthroat trout) (Salmo clarki).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=30962;
RN [1]
RW SEQUENCE FROM N.A.
RX MEDLINE=69382658; PubMed=2550657;
RA Thomas W.K.; Beckenbach A.T.;
RT "Variation in salmonid mitochondrial DNA: evolutionary constraints
RL and mechanisms of substitution."
J. Mol. Evol. 29:233-245(1989).
CC -1- FUNCTION: SUBUNIT I, II, AND III FORM THE FUNCTIONAL CORE OF
CC THE ENZYME COMPLEX.
CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
CC 4 FERRICYTOCHROME C.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.
DR HSSP: P00415; 10CC.
DR InterPro: IPR000298; CytC_oxdase_III.
DR Pfam: PF00510; COX3; 1.
DR PROSITE: PS50253; COX3; 1.
KW Oxidoreductase; Mitochondrion; Transmembrane.
SQ SEQUENCE 261 AA; 29705 MW; 434ECAA20954626E CRC64;

```

```

Query Match
Best Local Similarity 39.4%; Score 80; DB 1; Length 261;
Matches 17; Conservative 5; Mismatches 15; Indels 0; Gaps 0;
QY 1 IVHDKRRTAIGLAIGMIFTLQAYEYEVHT 37
Db 151 IMEGRKOTIALTLTLGFTFLQGMVEYEAFT 187

```

```

RESULT 8
ID COX3_ONCMY STANDARD; PRT; 261 AA.
AC P48172;
DT 01-FEB-1996 (Rel. 33, Created)

```

```

DT 01-FEB-1996 (Rel. 33, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE CYTOCHROME C OXIDASE POLYPEPTIDE III (EC 1.9.3.1).
GN COXIII OR COIII.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RW SEQUENCE FROM N.A.
RX TISSUE=Liver;
RX MEDLINE=96139027; PubMed=8587139;
RA Zardoya R., Garrido-Perterra A., Bautista J.M.;
RT "The complete nucleotide sequence of the mitochondrial DNA genome of
RT the rainbow trout, Oncorhynchus mykiss."
J. Mol. Evol. 41:942-951(1995).
CC -1- FUNCTION: SUBUNIT I, II, AND III FORM THE FUNCTIONAL CORE OF
CC THE ENZYME COMPLEX.
CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
CC 4 FERRICYTOCHROME C.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL: L29771; AAB03353.1;
DR HSSP: P00415; 10CC.
DR InterPro: IPR000298; CytC_oxdase_III.
DR Pfam: PF00510; COX3; 1.
DR PROSITE: PS50253; COX3; 1.
KW Oxidoreductase; Mitochondrion; Transmembrane.
SQ SEQUENCE 261 AA; 29675 MW; ADAF7AA21255D96E CRC64;

```

```

Query Match
Best Local Similarity 39.4%; Score 80; DB 1; Length 261;
Matches 17; Conservative 5; Mismatches 15; Indels 0; Gaps 0;
QY 1 IVHDKRRTAIGLAIGMIFTLQAYEYEVHT 37
Db 151 IMEGRKOTIALTLTLGFTFLQGMVEYEAFT 187

```

```

RESULT 9
ID COX3_ONCNE STANDARD; PRT; 261 AA.
AC P20683;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE CYTOCHROME C OXIDASE POLYPEPTIDE III (EC 1.9.3.1).
GN COXIII OR COIII.
OS Oncorhynchus masou (Sockeye salmon), and
OS Oncorhynchus masou (Cherry salmon) (Masu salmon).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8023; 8020;
RN [1]
RW SEQUENCE FROM N.A.
RX SPECIES=O. nerka;
RX MEDLINE=69382658; PubMed=2550657;
RA Thomas W.K.; Beckenbach A.T.;
RT "Variation in salmonid mitochondrial DNA: evolutionary constraints
RT and mechanisms of substitution."

```


OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 RX NCBI_TaxID=9913;
 RX [1]
 RX SEQUENCE FROM N.A.
 RX TISSUE-Heart;
 RX MEDLINE=83010260; PubMed=7120390;
 RX Anderson S., de Bruijn M.H.L., Coulson A.R., Eperon I.C., Sanger F.,
 RX Young I.G.;
 RX "Complete sequence of bovine mitochondrial DNA. Conserved features of
 RX the mammalian mitochondrial genome.";
 RX J. Mol. Biol. 156:683-717(1982).
 RX [2]
 RX X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
 RX MEDLINE=96216288; PubMed=8638158;
 RX Tsukihara T., Aoyama H., Yamashita E., Tomizaki T., Yamaguchi H.,
 RX Shinzawa-Itoh K., Nakashima R., Yaono R., Yoshikawa S.;
 RX "The whole structure of the 13-subunit oxidized cytochrome c oxidase
 RX at 2.8 A.";
 RX Science 272:1136-1144(1996).
 RX [3]
 RX X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
 RX TISSUE-Heart;
 RX MEDLINE=99190827; PubMed=10089392.
 RX Tomizaki T., Yamashita E., Yamaguchi H., Aoyama H., Tsukihara T.,
 RX Shinzawa-Itoh K., Nakashima R., Yaono R., Yoshikawa S.;
 RX "Structure analysis of bovine heart cytochrome c oxidase at 2.8 A
 RX resolution.";
 RX Acta Crystallogr. D 55:31-45(1999).
 RX [4]
 RX X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
 RX TISSUE-Heart;
 RX MEDLINE=20235645; PubMed=10771420;
 RX Fei M.J., Yamashita E., Inoue N., Yao M., Yamaguchi H., Tsukihara T.,
 RX Shinzawa-Itoh K., Nakashima R., Yoshikawa S.;
 RX "X-ray structure of azide-bound fully oxidized cytochrome c oxidase
 RX from bovine heart at 2.9 A resolution.";
 RX Acta Crystallogr. D 56:529-535(2000).
 RX [1]
 RX FUNCTION: SUBUNIT I, II, AND III FORM THE FUNCTIONAL CORE OF
 RX THE ENZYME COMPLEX.
 RX [2]
 RX CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
 RX 4 FERROCYTOCHROME C.
 RX [3]
 RX SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
 RX INNER MEMBRANE.
 RX [4]
 RX SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.
 RX
 RX This SWISS-PROT entry is copyright. It is produced through a collaboration
 RX between the Swiss Institute of Bioinformatics and the EMBL outstation -
 RX the European Bioinformatics Institute. There are no restrictions on its
 RX use by non-profit institutions as long as its content is in no way
 RX modified and this statement is not removed. Usage by and for commercial
 RX entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 RX or send an email to license@isb-sib.ch).
 RX -----
 RX EMBL: V00654; CAA24003.1; ALT_TERM.
 RX EMBL: J01394; AAB59274.1; -.
 RX PIR: A00483; OTB03.
 RX PDB: 1OCC; 07-DEC-96.
 RX PDB: 2OCC; 20CC; 26-MAY-98.
 RX PDB: 1OCC; 29-JUL-99.
 RX PDB: 1OCC; 22-JUL-99.
 RX PDB: 1OCC; 22-JUL-99.
 RX InterPro: IPR000298; Cytc_oxdse_III.
 RX Pfam: PF00510; COX3; 1.
 RX PROSITE: PS50253; COX3; 1.
 RX Oxioreductase; Mitochondrion; Transmembrane; 3D-structure;
 RX Inner membrane.
 RX KW
 RX TRANSMEM 1 15
 RX DOMAIN 16 34
 RX TRANSMEM 35 40
 RX DOMAIN 41 46
 RX TRANSMEM 41 66
 RX DOMAIN 67 72
 RX MITOCHONDRIAL MATRIX.
 RX MITOCHONDRIAL MATRIX.
 RX MITOCHONDRIAL MATRIX.

FT TRANSMEM 73 105
 FT DOMAIN 106 128
 FT TRANSMEM 129 152
 FT DOMAIN 153 155
 FT TRANSMEM 156 183
 FT DOMAIN 184 190
 FT TRANSMEM 191 223
 FT DOMAIN 224 232
 FT TRANSMEM 233 256
 FT DOMAIN 257 261
 FT SEQUENCE 261 AA; 29919 MW; 0347E76A9089B3F8 CRC64;
 SO
 Query Match 38.9%; Score 79; DB 1; Length 261;
 Best Local Similarity 51.4%; Pred. No. 0.0038;
 Matches 19; Conservative 2; Mismatches 16; Indels 0; Gaps 0;
 QY 1 IVHGDKRKTAIGAIAGIIGMTTCQAYEYVEIYHT 37
 Db 151 LMEDRRHMLQALFTITLTGVTTLQASEYVEAPFT 187
 RESULT 12
 COX3_CYPCA
 ID COX3_CYPCA STANDARD: PRT; 261 AA.
 AC P15952;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE CYTOCHROME C OXIDASE POLYPEPTIDE III (EC 1.9.3.1).
 GN COXIII OR COXII.
 OS Cyprinus carpio (Common carp).
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
 OC Cypriniformes; Cyprinidae; Cyprininae; Cyprinus.
 OX NCBI_TaxID=7962;
 RX [1]
 RX SEQUENCE FROM N.A.
 RX MEDLINE=90192097; PubMed=2156223;
 RX Huang C.J., Huang F.L., Chang Y.S., Tsai Y.J., Lo T.B.;
 RX "Nucleotide sequence of carp mitochondrial cytochrome c oxidase III.";
 RX Nucleic Acids Res. 18:1056-1056(1990).
 RX [2]
 RX SEQUENCE FROM N.A.
 RX MEDLINE=94223691; PubMed=8169959;
 RX Chang Y.S., Huang F.L., Lo T.B.;
 RX "The complete nucleotide sequence and gene organization of carp
 RX (Cyprinus carpio) mitochondrial genome.";
 RX J. Mol. Evol. 38:138-155(1994).
 RX [3]
 RX FUNCTION: SUBUNIT I, II, AND III FORM THE FUNCTIONAL CORE OF
 RX THE ENZYME COMPLEX.
 RX [4]
 RX CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
 RX 4 FERROCYTOCHROME C.
 RX [5]
 RX SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.
 RX
 RX This SWISS-PROT entry is copyright. It is produced through a collaboration
 RX between the Swiss Institute of Bioinformatics and the EMBL outstation -
 RX the European Bioinformatics Institute. There are no restrictions on its
 RX use by non-profit institutions as long as its content is in no way
 RX modified and this statement is not removed. Usage by and for commercial
 RX entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 RX or send an email to license@isb-sib.ch).
 RX -----
 RX EMBL: X17006; CAA34866.1; -.
 RX EMBL: X61010; CAA33341.1; -.
 RX PIR: S36010; OTCAS.
 RX HSSP: P00415; 1OCC.
 RX InterPro: IPR000298; Cytc_oxdse_III.
 RX Pfam: PF00510; COX3; 1.
 RX PROSITE: PS50253; COX3; 1.
 RX Oxioreductase; Mitochondrion; Transmembrane.
 RX CONFLICT 121 121
 RX I -> M (IN REF. 1).

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X72204; CAA51001.1; -
 DR PIR: S41826; S41826.
 DR HSSP: P00415; 10CC.
 DR InterPro: IPR000296; Cytc_oxdse_III.
 DR Pfam: PF00510; COX3; 1.
 DR PROSITE: PS50253; COX3; 1.
 KW Oxidoreductase; Mitochondrion; Transmembrane
 SQ SEQUENCE 261 AA; 29975 MW; EC1C7A4828260CA4 CRC64;

Query Match 37.9%; Score 77; DB 1; Length 261;
 Best Local Similarity 48.6%; Pred. No. 0.0068;
 Matches 18; Conservative 3; Mismatches 16; Indels 0; Gaps 0;
 QY 1 IVHGDRKKTAIGIAIGWITLQAYEYEVHT 37
 Db 151 LMEGNKRKHMQLFTIALGLFTLQASEYEAPET 187

Search completed: January 10, 2002, 02:03:12
 Job time: 204 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2002, 02:00:28 ; Search time 32.55 Seconds
(without alignments)
170.763 Million cell updates/sec

Title: US-09-712-768-6

Perfect score: 203
Sequence: 1 IVHGDRKKTALGALAIAGLWIFTLQAYEYEVHT 38

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP:archaea:*
2: SP:bacteria:*
3: SP:fungi:*
4: SP:human:*
5: SP_invertebrate:*
6: SP_mammal:*
7: SP_mhc:*
8: SP_organelle:*
9: SP_phage:*
10: SP_plant:*
11: SP_rodent:*
12: SP_virus:*
13: SP_vertbrate:*
14: SP_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	84	41.4	261	8 Q9MET0	Q9met0 physeter ca
2	84	41.4	267	8 Q9G7T0	Q9g7t0 sus scrofa
3	84	41.4	267	8 Q9G7S0	Q9g7s0 sus scrofa
4	84	41.4	278	8 Q9B290	Q9b290 sus scrofa
5	84	41.4	305	8 Q9T244	Q9t244 physophthor
6	83	40.9	261	8 Q9B629	Q9b629 anguilla ja
7	82	40.4	260	8 Q9GAV2	Q9gau2 rana sylvat
8	82	40.4	261	8 Q9B614	Q9b614 carassius a
9	81	39.9	261	8 Q9ZY39	Q9zy39 raja radiat
10	81	39.9	261	8 Q9XN34	Q9xn34 salvelinus
11	81	39.9	261	8 Q9XN26	Q9xn26 salvelinus
12	81	39.9	261	8 Q9G607	Q9g607 diplophos t
13	81	39.9	261	8 Q9G387	Q9g387 echinops te
14	81	39.9	261	8 Q9B263	Q9b263 isodon mac
15	80	39.4	261	8 Q9B263	Q9b263 oncorhynch
16	80	39.4	261	8 Q9B263	Q9b263 oncorhynch
17	80	39.4	261	8 Q9B603	Q9b603 pleocoglossu
18	79	38.9	261	8 Q9G6F4	Q9g6f4 polymixia j
19	79	38.9	290	2 Q9A304	Q9a304 caulobacter

20	78	38.4	261	8 Q9TEG5	Q9teg5 cavia porce
21	77	37.9	109	8 Q99872	Q99872 crotaphytus
22	77	37.9	109	8 Q99873	Q99873 crotaphytus
23	77	37.9	261	8 Q9MY4	Q9my4 brachydanio
24	77	37.9	261	8 Q9MEV4	Q9mev4 sciurus vul
25	76	37.4	262	8 Q9TBV9	Q9tbv9 acropora te
26	75	36.9	261	8 Q9XLO8	Q9xlo8 falco pereg
27	75	36.9	261	8 Q9MR64	Q9mr64 coregonus l
28	75	36.9	261	8 Q9G6R8	Q9g6r8 coregonus l
29	74	36.5	261	8 Q95418	Q95418 proteopterus
30	74	36.5	261	8 Q99600	Q99600 artibeus ja
31	74	36.5	261	8 Q9TBG8	Q9tbg8 turnix vari
32	74	36.5	261	8 Q9B6S8	Q9b6s8 eudromia el
33	74	36.5	262	8 Q9TA23	Q9ta23 loxodonta a
34	73	36.0	128	8 Q9T593	Q9t593 gambella w1
35	73	36.0	261	8 Q9TBG7	Q9tbg7 chaetura pe
36	73	36.0	261	8 Q9TBG5	Q9tbg5 chordelles
37	73	36.0	261	8 Q9MDB0	Q9mdb0 talpa europ
38	73	36.0	261	8 Q9G6T1	Q9g6t1 sardnops m
39	73	36.0	261	8 Q9B909	Q9b909 aulopus jap
40	72	35.5	76	8 Q48210	Q48210 drosophila
41	72	35.5	99	8 Q99876	Q99876 gambella w1
42	72	35.5	99	8 Q9T592	Q9t592 gambella w1
43	72	35.5	128	8 Q9T596	Q9t596 gambella w1
44	72	35.5	138	8 Q9XN56	Q9xn56 drosophila
45	72	35.5	138	8 Q9XK08	Q9xk08 drosophila

ALIGNMENTS

RESULT 1
ID Q9MET0 PRELIMINARY: PRT: 261 AA.
AC Q9MET0:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE CYTOCHROME OXIDASE SUBUNIT III.
GN COIT1.
OS Physeter catodon (Sperm whale) (Physeter macrocephalus).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Odontoceti;
OC Physeteridae; Physeter.
OX NCBI_TaxID=9755;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20296818; PubMed=10835487;
RA Arnason U., Gulberg A., Gretaerdotir S., Ursing B., Janke A.;
RT "The mitochondrial genome of the sperm whale and a new molecular
RT reference for estimating eutherian divergence dates.";
RL J. Mol. Evol. 50:569-578(2000).
DR EMBL: A0277029; CAB96276.1; -;
DR InterPro: IPR000298; CYC-oxide_III.
DR Pfam: PF00510; COX3; 1.
DR PROSITE: PS50253; COX3; 1.
KW Mitochondrion.
SQ SEQUENCE 261 AA; 29907 MW; B7010A5A0AF2F28B CRC64;

Query Match 41.4%; Score 84; DB 8; Length 261;
Best local Similarity 51.4%; Pred. No. 0.0023;
Matches 19; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

QY 1 IVHGDRKKTALGALAIAGLWIFTLQAYEYEVHT 37
Db 151 IHWGDRKKTALGALAIAGLWIFTLQAYEYEVHT 187
RESULT 2
Q9G7T0 PRELIMINARY: PRT: 267 AA.

AC Q9G7T0: 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE COI1.
 GN COI1.
 OS Sus scrofa (Pig).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kijas J.M.H., Andersson L.;
 RT "A phylogenetic study of the origin of the domestic pig estimated from
 the near complete mtDNA genome."
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF304201; AAG28196.1;
 DR InterPro; IPR000298; Cytc_oxdse_III.
 DR Pfam; PF00510; COX3; 1.
 DR PROSITE; PS50253; COX3; 1.
 KW Mitochondrion.
 SQ SEQUENCE 267 AA; 30462 MW; 95C558C54F6C0E3F CRC64;

Query Match 41.4%; Score 84; DB 8; Length 267;
 Best Local Similarity 51.4%; Pred. No. 0.0023;
 Matches 19; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

OY 1 IVHGDRKKTALGALAIAGLGMIFLTCQAYEYEVHT 37
 DB 151 LMEDGRKHMIALSTITALGVFTLLQASFEYEAFT 187

RESULT 3
 ID Q9G7S0: PRELIMINARY; PRT; 267 AA.
 AC Q9G7S0;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE COI1.
 GN COI1.
 OS Sus scrofa (Pig).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kijas J.M.H., Andersson L.;
 RT "A phylogenetic study of the origin of the domestic pig estimated from
 the near complete mtDNA genome."
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF304203; AAG28222.1;
 DR InterPro; IPR000298; Cytc_oxdse_III.
 DR Pfam; PF00510; COX3; 1.
 DR PROSITE; PS50253; COX3; 1.
 KW Mitochondrion.
 SQ SEQUENCE 267 AA; 30454 MW; 471AED9C294867E9 CRC64;

Query Match 41.4%; Score 84; DB 8; Length 267;
 Best Local Similarity 51.4%; Pred. No. 0.0023;
 Matches 19; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

OY 1 IVHGDRKKTALGALAIAGLGMIFLTCQAYEYEVHT 37
 DB 151 LMEDGRKHMIALSTITALGVFTLLQASFEYEAFT 187

RESULT 4
 Q9G290

ID Q9G290: PRELIMINARY; PRT; 278 AA.
 AC Q9G290;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE COI1.
 GN COI1.
 OS Sus scrofa (Pig).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kijas J.M.H., Andersson L.;
 RT "A phylogenetic study of the origin of the domestic pig estimated from
 the near complete mtDNA genome."
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF304202; AAG28209.1;
 DR EMBL; AF304200; AAG28183.1;
 DR InterPro; IPR000298; Cytc_oxdse_III.
 DR Pfam; PF00510; COX3; 1.
 DR PROSITE; PS50253; COX3; 1.
 KW Mitochondrion.
 SQ SEQUENCE 278 AA; 31927 MW; 25A60814452E5A2B CRC64;

Query Match 41.4%; Score 84; DB 8; Length 278;
 Best Local Similarity 51.4%; Pred. No. 0.0024;
 Matches 19; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

OY 1 IVHGDRKKTALGALAIAGLGMIFLTCQAYEYEVHT 37
 DB 151 LMEDGRKHMIALSTITALGVFTLLQASFEYEAFT 187

RESULT 5
 ID Q9T244: PRELIMINARY; PRT; 305 AA.
 AC Q9T244;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE CYTOCHROME C OXIDASE SUBUNIT 3 (EC 1.9.3.1).
 GN COX3.
 OS Phytophthora infestans (Potato late blight fungus).
 OG Mitochondrion.
 OC Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
 OC Phytophthora.
 OX NCBI_TaxID=4787;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WEST VIRGINIA 4;
 RA Lang B.F., Forget L.;
 RT "The mitochondrial genome of phytophthora infestans."
 RL (In) O'Brien S.J. (eds.);
 RL Genetic Maps, pp.3:133-135, Cold Spring Harbor Laboratory Press,
 NY (1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WEST VIRGINIA 4;
 RA Paquin B., Roewer I., Wang Z., Lang B.F.;
 RT "A robust fungal phylogeny using the mitochondrially encoded mads
 protein sequence."
 RL Can. J. Bot. 73:5180-5185(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WEST VIRGINIA 4;
 RA MEDLINE=97306266; PubMed=9162109;
 RA Paquin B., Laforest M.J., Forget L., Roewer I., Wang Z., Longcore J.,
 RA Lang B.F.;
 RT "The fungal mitochondrial genome project: evolution of fungal
 mitochondrial genomes and their gene expression."

[REDACTED]

100

100

RX MEDLINE=99162577; Pubmed=10051614;
 RA Rasmussen A.-S.; Arnason U.;
 RT "Molecular studies suggest that cartilaginous fishes have a terminal
 position in the piscine tree."
 RL Proc. Natl. Acad. Sci. U.S.A. 96:2177-2182(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RA Rasmussen A.-S.;
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF106038; AAD19935.1; -
 DR HSSP: P18402; 1FFT
 DR InterPro: IPR000298; Cytc_oxdse_III.
 DR Pfam: PF00510; COX3; 1.
 DR PROSITE: PS50253; COX3; 1.
 DR Mitochondrion.
 SQ SEQUENCE 261 AA; 29600 MW; 6A9B1A6DC779F788 CRC64;

Query Match 39.9%; Score 81; DB 8; Length 261;
 Best Local Similarity 45.9%; Pred. No. 0.0057;
 Matches 17; Conservative 5; Mismatches 15; Indels 0; Gaps 0;
 OY 1 IVHGRRKTAIGLAIAIGLWIFLCOAYEYEVHT 37
 Db 151 IMEGRRKQTVALTLTLGLGFYFLQAMEYEAAPT 187

RESULT 10
 O9XN34 PRELIMINARY; PRT; 261 AA.
 AC O9XN34.
 DT 01-NOV-1999 (TReMBLrel. 12, Created)
 DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
 DE CYTOCHROME OXIDASE SUBUNIT 3.
 GN COIII.
 OS Salvelinus fontinalis (Brook trout) (Brook char).
 OG Mitochondrion.
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 CC Protacanthopterygii; Salmoniformes; Salmonidae; Salvelinus.
 OX NCBI_TaxID=8036;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Doiron S., Blier P.U., Bernatchez L.;
 RT "A comparative analysis of complete sequence of mitochondrial genome
 between brook char (Salvelinus fontinalis) and arctic char (S.
 alpinus)."
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF154850; AAD41378.1; -
 DR HSSP: P18402; 1FFT
 DR InterPro: IPR000298; Cytc_oxdse_III.
 DR Pfam: PF00510; COX3; 1.
 DR PROSITE: PS50253; COX3; 1.
 DR Mitochondrion.
 SQ SEQUENCE 261 AA; 29721 MW; 9EAA4C788FE00316 CRC64;

Query Match 39.9%; Score 81; DB 8; Length 261;
 Best Local Similarity 45.9%; Pred. No. 0.0057;
 Matches 17; Conservative 5; Mismatches 15; Indels 0; Gaps 0;
 OY 1 IVHGRRKTAIGLAIAIGLWIFLCOAYEYEVHT 37
 Db 151 IMEGRRKQTVALTLTLGLGFYFLQAMEYEAAPT 187

RESULT 11
 O9XN26 PRELIMINARY; PRT; 261 AA.
 AC O9XN26.
 DT 01-NOV-1999 (TReMBLrel. 12, Created)

DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
 DE CYTOCHROME OXIDASE SUBUNIT 3.
 GN COIII.
 OS Salvelinus alpinus (Arctic char).
 OG Mitochondrion.
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 CC Protacanthopterygii; Salmoniformes; Salmonidae; Salvelinus.
 OX NCBI_TaxID=8036;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Doiron S., Blier P.U., Bernatchez L.;
 RT "A comparative analysis of complete sequence of mitochondrial genome
 between brook char (Salvelinus fontinalis) and arctic char (S.
 alpinus)."
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF154851; AAD41391.1; -
 DR HSSP: P18402; 1FFT
 DR InterPro: IPR000298; Cytc_oxdse_III.
 DR Pfam: PF00510; COX3; 1.
 DR PROSITE: PS50253; COX3; 1.
 DR Mitochondrion.
 SQ SEQUENCE 261 AA; 29751 MW; A96564CE40C15BA9 CRC64;

Query Match 39.9%; Score 81; DB 8; Length 261;
 Best Local Similarity 45.9%; Pred. No. 0.0057;
 Matches 17; Conservative 5; Mismatches 15; Indels 0; Gaps 0;
 OY 1 IVHGRRKTAIGLAIAIGLWIFLCOAYEYEVHT 37
 Db 151 IMEGRRKQTVALTLTLGLGFYFLQAMEYEAAPT 187

RESULT 12
 O9G607 PRELIMINARY; PRT; 261 AA.
 AC O9G607.
 DT 01-MAR-2001 (TReMBLrel. 16, Created)
 DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
 DE CYTOCHROME C OXIDASE SUBUNIT III.
 GN COIII.
 OS Diplaphos taenia.
 OG Mitochondrion.
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 CC Stomliiformes; Gonostomatidae; Diplaphos.
 OX NCBI_TaxID=91975;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Miya M., Nishida M.;
 RT "Setting a new stage for the teleostean molecular systematics: A
 submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB034825; BAB20726.1; -
 DR InterPro: IPR000298; Cytc_oxdse_III.
 DR Pfam: PF00510; COX3; 1.
 DR PROSITE: PS50253; COX3; 1.
 DR Mitochondrion.
 SQ SEQUENCE 261 AA; 29585 MW; 694856C933298B77 CRC64;

Query Match 39.9%; Score 81; DB 8; Length 261;
 Best Local Similarity 48.6%; Pred. No. 0.0057;
 Matches 18; Conservative 5; Mismatches 14; Indels 0; Gaps 0;
 OY 1 IVHGRRKTAIGLAIAIGLWIFLCOAYEYEVHT 37
 Db 151 IMEGRRKQTVALTLTLGLGFYFLQAMEYEAAPT 187

RESULT 13
 ID 09G387 PRELIMINARY; PRT: 261 AA.
 AC 09G387;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, last annotation update)
 DE CYTOCHROME OXIDASE SUBUNIT III.
 GN COII.
 OS Echinops telfairi (Lesser hedgehog tenrec).
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Insectivora; Tenrecidae; Echinops.
 OX NCBI_TaxID=9371;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Mouchaty S.K., Gullberg A., Janke A., Arnason U.;
 RT "Phylogenetic position of the tenrecs (Mammalia: Tenrecidae) of
 RT Madagascar based on analysis of the complete mitochondrial genome
 RT sequence of Echinops telfairi.";
 RL Zool. Scr. 29:307-317(2000).
 DR EMBL: AJ400734; CAC19398.1; -
 DR InterPro: IPR000298; Cyto_oxdse_III.
 DR Pfam: PF00510; COX3; 1.
 DR PROSITE: PS50253; COX3; 1.
 KW Mitochondrion.
 SO SEQUENCE 261 AA; 29773 MW; B66DEB567687DD58 CRC64;

Query Match 39.9%; Score 81; DB 8; Length 261;
 Best Local Similarity 48.6%; Pred. No. 0.0057;
 Matches 18; Conservative 5; Mismatches 14; Indels 0; Gaps 0;
 Oy 1 IVHGRRKTAIGLAIAGLWIFLLQAYEYEVHT 37
 Db 151 IMEGNRKQMIQALSTIALGLYFTLLQAMEYEAFT 187
 RESULT 14
 ID 09B263 PRELIMINARY; PRT: 261 AA.
 AC 09B263;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, last annotation update)
 DE CYTOCHROME OXIDASE SUBUNIT III.
 GN COII.
 OS Isodon macrourus (Short-nosed bandicoot).
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Metatheria; Peramellemorphia; Peramelidae; Isodon.
 OX NCBI_TaxID=37698;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Phillips M.J., Lin Y.-H., Harrison G.L., Penny D.;
 RT "Mitochondrial Genomes of a Bandicoot and a Brush-tail Possum Confirm
 RT the Monophyly of Australidelphian Marsupials.";
 RL Proc. R. Soc. Lond. B, Biol. Sci. 0:0-0(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Lin Y.-H., Phillips M.J.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF358864; AAK38287.1; -
 KW Mitochondrion.
 SO SEQUENCE 261 AA; 29966 MW; A108114486943414 CRC64;

Db 151 IMEGNRKQMIQALSTIALGLYFTLLQAMEYEAFT 187
 RESULT 15
 ID 003378 PRELIMINARY; PRT: 261 AA.
 AC 003378;
 DT 01-JUL-1997 (TREMBlrel. 04, Created)
 DT 01-JUL-1997 (TREMBlrel. 04, last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, last annotation update)
 DE CYTOCHROME OXIDASE SUBUNIT 3.
 OS Oncorhynchus keta (Chum salmon).
 OC Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OX NCBI_TaxID=8018;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN-TOKORO RIVER, HOKKAIDO, JAPAN;
 RC MEDLINE=97159563; PubMed=9007022;
 RA Oohara I., Sawano K., Okazaki T.;
 RT "Mitochondrial DNA sequence analysis of the masu salmon--phylogeny in
 RT the genus Oncorhynchus.";
 RL Mol. Phylogenet. Evol. 7:71-78(1997).
 DR EMBL: D84147; BAA20154.1; -
 DR HSSP: P18402; 1PFT
 DR InterPro: IPR000298; Cyto_oxdse_III.
 DR Pfam: PF00510; COX3; 1.
 DR PROSITE: PS50253; COX3; 1.
 KW Mitochondrion.
 SO SEQUENCE 261 AA; 29691 MW; 4EF55462ACD66DF CRC64;

Query Match 39.4%; Score 80; DB 8; Length 261;
 Best Local Similarity 45.9%; Pred. No. 0.0077;
 Matches 17; Conservative 5; Mismatches 15; Indels 0; Gaps 0;
 Oy 1 IVHGRRKTAIGLAIAGLWIFLLQAYEYEVHT 37
 Db 151 IMEGNRKQMIQALSTIALGLYFTLLQAMEYEAFT 187

Search completed: January 10, 2002, 02:04:05
 Job time: 217 sec

Query Match 39.9%; Score 81; DB 8; Length 261;
 Best Local Similarity 45.9%; Pred. No. 0.0057;
 Matches 17; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

Oy 1 IVHGRRKTAIGLAIAGLWIFLLQAYEYEVHT 37
 Db 151 IMEGNRKQMIQALSTIALGLYFTLLQAMEYEAFT 187

us-09-712-768-6.rni

seq. name	/cg21_6/prodata/2/lna/6B_COMB.seq	seq:US-09-020-956-6
/cg21_6/prodata/2/lna/5B_COMB.seq <td>US-08-633-770A-12 +</td> <td>47.50 99.03 724.75</td>	US-08-633-770A-12 +	47.50 99.03 724.75
/cg21_6/prodata/2/lna/6A_COMB.seq <td>US-09-128-155-17 +</td> <td>47.50 62.89 7.1e+04</td>	US-09-128-155-17 +	47.50 62.89 7.1e+04
/cg21_6/prodata/2/lna/5B_COMB.seq <td>US-09-105-989-6 -</td> <td>47.00 112.94 121.80</td>	US-09-105-989-6 -	47.00 112.94 121.80
/cg21_6/prodata/2/lna/6A_COMB.seq <td>US-09-138-922-6 -</td> <td>47.00 112.94 121.80</td>	US-09-138-922-6 -	47.00 112.94 121.80
/cg21_6/prodata/2/lna/6A_COMB.seq <td>US-09-079-981-3 -</td> <td>47.00 105.99 296.95</td>	US-09-079-981-3 -	47.00 105.99 296.95
seq. name	/cg21_6/prodata/2/lna/6B_COMB.seq	seq:US-09-020-956-6

```

seq_name: /cogn2_6/plodata/2/lna/6B-COMB.seq:US-09-020-956-6
seq_documentation_block:
; Sequence 6, Application US/09020956
; Patent No. 6261562
;
; GENERAL INFORMATION:
;
; APPLICANT: Xu, Jiangchun
;
; APPLICANT: Dillin, Davin C.
;
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS
;
; NUMBER OF SEQUENCES: 178
;
; CORRESPONDENCE ADDRESS:
;
; ADDRESSEE: SEED and BERRY LLP
;
; STREET: 6300 Columbia Center, 701 Fifth Avenue
;
; CITY: Seattle
;
; STATE: WA
;
; COUNTRY: USA
;

```

```

1  ZLP: 96104
2  COMPUTER READABLE FORM:
3  MEDIUM TYPE: Floppy disk
4  COMPUTER: IBM PC compatible
5  OPERATING SYSTEM: PC-DOS/MS-DOS
6  SOFTWARE: PatentIn Release #1.0, Version #1.30
7  CURRENT APPLICATION DATA:
8  APPLICATION NUMBER: US/09/020,956

```

```

; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.427C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 818 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-020-956-6

alignment_scores:
    Quality: 62.00      Length: 33
    Ratio: 2.583        Gaps: 0
Percent Similarity: 72.727    Percent Identity: 42.424

alignment_block:
US-09-712-768-6 x US-09-020-956-6/rev ..

Align seq 1/1 to reverse of: US-09-020-956-6 from: 1 to: 818

1 11evaihsclysaphargylslysthralaalleglyleualaalleala1 17
      :::::::::::::::::::::|||||
354 CTATACGAAACAAACGAAACCAATTAATTCACGACCTGCTTATTACAT 305

17 eglvleuglytirpleptherleucysginalatyrglutyrtyrglu 33
      ||||||| ||||||| |||||||
304 TTTCCTCGGTCCTATTATTACCTCTACACAGCCTCAGAGTACTTCGAG 256

seq_name: /cgu2_6/plodata/2/lna/6B_COMB.seq:US-09-030-607-6

seq_documentation_block:
; Sequence 6, Application US/09030607
; Patent No. 6262245
; GENERAL INFORMATION:

```

APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS FO
NUMBER OF SEQUENCES: 224
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/030,607
FILING DATE: 25-FEB-1998
CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.427C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 682-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 818 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: CDNA
US-09-030-607-6

alignment_scores:

Quality: 62.00 Length: 33
Ratio: 2.583 Gaps: 0
Percent Similarity: 72.727 Percent Identity: 42.424

alignment_block:

US-09-712-768-6 x US-09-030-607-6/rev ..

Align seg 1/1 to reverse of: US-09-030-607-6 from: 1 to: 818

```
1 llevalhlsiglyasparglysthrAlaileglyleuAlaileAla1 17
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
354 CTATAGAAACACGAAACCAATATTCAGACACTGCTTATATACAT 305
17 eglyleuglyTrrpIlephethrleucysglnAlaTyrgIuTyrglu 33
304 TTTACTGGGCTCTATTATTACCTCTCTACAGCCTCAGAGTACTTCGAG 256
```

seq_name: /cgn2_6/plodata/2/lna/5A_COMB.seq:US-08-219-842-3

seq_documentation_block:

Sequence 3, Application US/08219842
Patent No. 5563323
GENERAL INFORMATION:
APPLICANT: Parker, W. D.
APPLICANT: Herrnstadt, Corina
TITLE OF INVENTION: Diagnostic and Therapeutic Compositions
NUMBER OF SEQUENCES: for Alzheimer's Disease
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/219,842
FILING DATE: 30-MAR-1994
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-AG 9504
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 856 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

US-08-219-842-3

alignment_scores:

Quality: 62.00 Length: 33
Ratio: 2.583 Gaps: 0
Percent Similarity: 72.727 Percent Identity: 42.424

alignment_block:

US-09-712-768-6 x US-08-219-842-3 ..

Align seg 1/1 to: US-08-219-842-3 from: 1 to: 856

```
1 llevalhlsiglyasparglysthrAlaileglyleuAlaileAla1 17
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
521 CTATAGAAACACGAAACCAATATTCAGACACTGCTTATATACAT 570
17 eglyleuglyTrrpIlephethrleucysglnAlaTyrgIuTyrglu 33
571 TTTACTGGGCTCTATTATTACCTCTCTACAGCCTCAGAGTACTTCGAG 619
```

seq_name: /cgn2_6/plodata/2/lna/5A_COMB.seq:US-08-451-096-3

seq_documentation_block:

Sequence 3, Application US/08451096
Patent No. 5760205
GENERAL INFORMATION:
APPLICANT: Parker, W. D.
APPLICANT: Herrnstadt, Corina
TITLE OF INVENTION: Diagnostic and Therapeutic Compositions
NUMBER OF SEQUENCES: for Alzheimer's Disease
CORRESPONDENCE ADDRESS:
ADDRESSEE: Campbell and Flores
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/451,096
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/219,842
FILING DATE: 30-MAR-1994

us-09-712-768-6.rni

ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-AG 9504
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 355-8949
INFORMATION FOR SEO ID NO.: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 856 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

alignment_scores:	62.00	Length:	33
Quality:	2.583	Gaps:	0
Ratio:	72.727	Percent Identity:	42.424

```
alignment_block:
ms-09-712-768-6 x US-08-451-096-3 .
```

Align seg 1/1 to: US-08-451-096-3 from: 1 to: 856

1 llevalhlsiglyasparglyslysthrlealleeglyleualleleall 17
 521 cfnatgaaaacacccagcaaccaaataattcagacgctctattacat 578
 17 eglyleuglyttrpilesphethrleucysglinalatryglutryrglu 33
 571 tttactggtctctatttttaccctctacacacccctcagagacttgag 619
 572 tttactggtctctatttttaccctctacacacccctcagagacttgag 619

```
seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-810-599-82
```

```
seq_documentation_block:
  sequence 82, Application US/08810599
```

GENERAL INFORMATION: PARKER, W. Davis
APPLICANT: HERRNSTADT, Corinna
APPLICANT: GHOSH, Soumitra S.
APPLICANT: PARY, Eoin
TITLE OF INVENTION: Methods for Detecting Mitochondrial Mutations
TITLE OF INVENTION: Diagnostic for Alzheimer's Disease and Methods for Determining
TITLE OF INVENTION: of Mitochondrial Nucleic Acid
NUMBER OF SEQUENCES: 82
CORRESPONDENCE ADDRESS: Kenyon & Kenyon
ADDRESSEE: Kenyon & Kenyon, Avenue, N.W., Suite 600

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

APPLICATION NUMBER: US 08/413,740
FILING DATE: 30 Mar 1995
APPLICATION NUMBER: US 08/410,658
FILING DATE: 24 MARCH 1995
APPLICATION NUMBER: US 08/397,808
FILING DATE: 3 Mar 1995
APPLICATION NUMBER: US 08/219,843
FILING DATE: 30 MARCH 1994
ATTORNEY/AGENT INFORMATION:
NAME: Toftennett, Judith L.
REGISTRATION NUMBER: 39,048
REFERENCE/DOCKET NUMBER: 2105/17
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-429-1776
TELEFAX: 202-429-0796
INFORMATION FOR SEQ ID NO: 82:
SEQUENCE CHARACTERISTICS:
LENGTH: 926 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-810-599-82

alignment_scores:	length:	33
quality:	gaps:	0
ratio:	percent identity:	42.424
percent similarity:		72.727

alignment_block: US-08-810-599-82

Align seg 1/1 to: US-08-810-599-82 from: 1 to: 926

[illegible]

```
seq_name: /cgn2_6/ptodata/2/ina/6B_COMB.seq:05-09-09/...
```

```
seq_documentation_block:
  sequence 2, Application US/09097889
```

GENERAL INFORMATION:
APPLICANT: Herrinstad, Corina
APPLICANT: Ghosh, Soumitra S.
APPLICANT: Davis, Robert E.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING
AGENTS THAT QUANTITATIVELY ALTER DETECTABLE
TITLE OF INVENTION: MITOCHONDRIAL DNA RATIOS
TITLE OF INVENTION: MITOCHONDRIAL DNA
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS: 10000 110

CORRESPONDENCE: SEED and BERRY LLP
 ADDRESS: 6300 Columbia Center, 701 Fifth Avenue
 STREET: 6300 Columbia Center, 701 Fifth Avenue
 CITY: Seattle
 STATE: Washington
 COUNTRY: USA
 ZIP: 98104
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.33
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/097, 889
 FILING DATE: 15-JUN-1998

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: 1

Thu Jan 10 08:15:48 2002

us-09-712-768-6.rn1

Page 5

```

; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAG1453UP
US-08-998-416-684

alignment_scores:
  Quality: 56.00      Length: 31
  Ratio: 2.667        Gaps: 0
  Percent Similarity: 67.742      Percent Identity: 41.935

alignment_block:
US-09-712-768-6 x US-08-998-416-684 ..

Align seg 1/1 to: US-08-998-416-684 from: 1 to: 732

1 11evah1hsglyasparglysthrAla11leGlyLeuAla11leAl1 17
   :::::::::::::::::::::  ::::  ::::  ::::  ::::
460 TTAATTGAAGTAATAGAAACATGCTTTATCAAGTTTACTTACTTT 509
17 ecglyLeuGlyTTPlePheThrLeuGlyGlnAla1aTyrgluTyr 31
   ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::
510 CTGATTAAATTGTTACTTTGTTATGTCAATATATGAAATAT 552

seq_name: /cgn2_6/plodata/2/lna/6B_COMB.seq:US-08-998-416-306

seq_documentation_block:
; Sequence 306, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jürgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Reibschung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSTYPII
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; City: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/998,416
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Weigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8689
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 306:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 847 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
```

```

; ORIGINAL SOURCE:
; ORGANISM: PAG1250UP
US-08-998-416-306

alignment_scores:
  Quality: 56.00      Length: 31
  Ratio: 2.667        Gaps: 0
  Percent Similarity: 67.742      Percent Identity: 41.935

alignment_block:
US-09-712-768-6 x US-08-998-416-306 ..

Align seg 1/1 to: US-08-998-416-306 from: 1 to: 847

1 11evah1hsglyasparglysthrAla11leGlyLeuAla11leAl1 17
   :::::::::::::::::::::  ::::  ::::  ::::  ::::
459 TTAATTGAAGTAATAGAAACATGCTTTATCAAGTTTACTTACTTT 508
17 ecglyLeuGlyTTPlePheThrLeuGlyGlnAla1aTyrgluTyr 31
   ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::
509 CTGATTAAATTGTTACTTTGTTATGTCAATATATGAAATAT 551

seq_name: /cgn2_6/plodata/2/lna/6A_COMB.seq:US-09-356-952-12

seq_documentation_block:
; Sequence 12, Application US/09356952
; Patent No. 6117663
; GENERAL INFORMATION:
; APPLICANT: Borlack-Sjodin, Ann
; APPLICANT: Margalit, S. M.
; APPLICANT: Bor-Sog1, Dafna
; APPLICANT: Cole, Philip
; APPLICANT: Kuriyan, John
; TITLE OF INVENTION: A CRYSTAL OF A RAS-SOS COMPLEX AND METHODS OF USE
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 600-1-228N
; CURRENT APPLICATION NUMBER: US/09/356,952
; CURRENT FILING DATE: 1999-07-19
; EARLIER APPLICATION NUMBER: 60/093,631
; EARLIER FILING DATE: 1998-07-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 12
; LENGTH: 43676
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
; US-09-356-952-12

alignment_scores:
  Quality: 54.00      Length: 32
  Ratio: 2.842        Gaps: 1
  Percent Similarity: 59.375      Percent Identity: 46.875

alignment_block:
US-09-712-768-6 x US-09-356-952-12/rev ..

Align seg 1/1 to reverse of: US-09-356-952-12 from: 1 to: 43676

1 11leGlyLeuAla11leAla11leGlyLeuGlyTTPle.....Ph 23
   ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::
497 ATTAGCTACAGAAACATAGATGATGACATCAAGACACTTT 448
23 ethrLeuGlyGlnAla1aTyrgluTyrGlu11leAla11leAla 38
   ::::  ::::  ::::  ::::  ::::  ::::  ::::  ::::
447 TACGTTGCTGCTGCTTATTTATGTAATCAATCATGACGACG 402

seq_name: /cgn2_6/plodata/2/lna/PCBUS_COMB.seq:PCT-US91-09422-20

seq_documentation_block:
; Sequence 20, Application PC/TUS9109422
; GENERAL INFORMATION:
```

APPLICANT: Mulvihill, Eileen R.
 APPLICANT: Hagen, Frederick S.
 APPLICANT: Hounamed, Khaled M.
 APPLICANT: Almers, Wolfhard
 TITLE OF INVENTION: G PROTEIN-COUPLED GLUTAMATE RECEPTORS
 NUMBER OF SEQUENCES: 33
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Townsend and Townsend
 STREET: One Market Plaza, Stewart Street Tower
 CITY: San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94105-1492
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC Compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US91/09422
 FILING DATE: 19911212
 CLASSIFICATION: 435
 Prior APPLICATION DATA:
 APPLICATION NUMBER: US 07/672,007
 FILING DATE: 18-MAR-1991
 Prior APPLICATION DATA:
 APPLICATION NUMBER: US 07/648,481
 FILING DATE: 30-JAN-1991
 Prior APPLICATION DATA:
 APPLICATION NUMBER: US 07/626,806
 FILING DATE: 12-DEC-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Parmelee, Steven W.
 REGISTRATION NUMBER: 31,990
 REFERENCE/DOCKET NUMBER: 19952-6PC
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (206) 467-9600
 TELEFAX: (415) 543-5043
 INFORMATION FOR SEQ ID NO: 20:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2426 base pairs
 TYPE: NUCLEIC ACID
 STRANDEDNESS: Single
 TOPOLOGY: Linear
 MOLECULE TYPE: cDNA
 IMMEDIATE SOURCE:
 CLONE: SR13

alignment_scores:	
Quality:	50.00
Ratio:	2.941
Percent Similarity:	60.714
	Length:
	Gaps:
	1
	Percent Identity:
	39.286

alignment_block:

Align seg 1/1 to: PCT-US91-09422-20

[illegible]

```
seq_name: /cgn2_6/prodata/2/1na/PCTUS_COMB.seq:PCT-US91-09422-18
seq_documentation_block:
; Sequence 18. Application PC/TUS9109422
; GENERAL INFORMATION:
```

APPLICANT: Mulvihill, Eileen R.
 APPLICANT: Hagen, Frederick S.
 APPLICANT: Houmed, Khalid M.
 APPLICANT: Almers, Wolfgang
 TITLE OF INVENTION: G PROTEIN-COUPLED GLUTAMATE RECEPTORS
 NUMBER OF SEQUENCES: 33
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Townsend and Townsend
 STREET: One Market Plaza, Stewart Street Tower
 CITY: San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94105-1492
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US91/09422
 FILING DATE: 19911212
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/672,007
 FILING DATE: 18-MAR-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/648,481
 FILING DATE: 30-JAN-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/626,806
 FILING DATE: 12-DEC-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: Parmelee, Steven W.
 REGISTRATION NUMBER: 31,990
 REFERENCE/DOCKET NUMBER: 13952-6PC
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (206) 467-9600
 TELEFAX: (415) 543-5043
 INFORMATION FOR SEQ ID NO: 18:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 4095 base pairs
 TYPE: NUCLEIC ACID
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 IMMEDIATE SOURCE:
 CLONE: SN30
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 463..3198
 CDT-US91-09422-18

alignment_scores:

Quality:	50.00	Length:	28
Ratio:	2.941	Gaps:	1
Percent Similarity:	60.714		

alignment_block:

Align seg 1/1 to: PCN-US91-09422-18 From: 1 to: 4095

[illegible]


```

seq_name: /cgn2_6/ptodata/2/lna/6A.COMB.seq:US-09-215-131-1
    seq_documentation_block:
        ; Sequence 1, Application US/09215131
        ; Patent No. 6030834
        ; GENERAL INFORMATION:
        APPLICANT: Pot, David
        TITLE OF INVENTION: IKK Beta Regulates Transcription Factors
        FILE REFERENCE: 1449.002
        CURRENT APPLICATION NUMBER: US/09/215,131
        CURRENT FILING DATE: 1998-12-18
        NUMBER OF SEQ ID NOS: 3
        SOFTWARE: FastSeq for Windows Version 3.0
        SEQ ID NO 1
            LENGTH: 3966
            TYPE: DNA
            ORGANISM: human
        US-09-215-131-1

alignment_scores:
    Quality:      49.50          Length:      30
    Ratio:         2.605           Gaps:       2
    Percent Similarity: 63.333     Percent Identity: 36.667

alignment_block:
    US-09-712-768-6 x US-09-215-131-1 ..

Align seg 1/1   to: US-09-215-131-1   from: 1   to: 3966

                3 HHSGLYASPARGLYSISRHRALAIL.....GlyLe 13
                |||||              || ::::~:|||||
3483 CATGGAGATGAATGAAAGCATTCGTGCTCAAGTAAGTTAGGAGACGATT 3532
                    13 uAlAIleAlAIleGIyLeuGlYTyrPIIePherhrlEuCs 26
                        : ::::::::::::::| | | | | | | | | | |
                        3533 TCATGTTTCACTTGCTTGG...TGGAGATTCCACTATGC 3569

seq_name: /cgn2_6/ptodata/2/lna/6A.COMB.seq:US-09-222-734-1
    seq_documentation_block:
        ; Sequence 1, Application US/09222734A
        ; Patent No. 6077701
        ; GENERAL INFORMATION:
        APPLICANT: Chu, Keling
        TITLE OF INVENTION: IKK-beta Regulates Transcription Factors
        FILE REFERENCE: 12441.78080
        CURRENT APPLICATION NUMBER: US/09/222,734A
        CURRENT FILING DATE: 1998-12-29
        EARLIER APPLICATION NUMBER: 09/215,131
        EARLIER FILING DATE: 1998-12-18
        EARLIER APPLICATION NUMBER: 60/068,954
        EARLIER FILING DATE: 1997-12-30
        NUMBER OF SEQ ID NOS: 3
        SOFTWARE: Patentin Ver. 2.0
        SEQ ID NO 1
            LENGTH: 3966
            TYPE: DNA
            ORGANISM: Homo sapiens
        US-09-222-734-1

alignment_scores:
    Quality:      49.50          Length:      30
    Ratio:         2.605           Gaps:       2
    Percent Similarity: 63.333     Percent Identity: 36.667

alignment_block:
```

```

US-09-712-768-6 x US-09-222-734-1 ..
Align seg 1/1 to: US-09-222-734-1 from: 1 to: 3966

3 HisGlyAspArgLysLysThrAlaIle.....GlyLe 13
|||||  |||  ::::
3483 CATGAGATAGATGATGACATCTGTGCTCAGTAACTTTAGGAGACTAT 3532
13 uAlaIleAlaIleGlyLeuGlyTrpIlePheThrLeuCys 26
: :::::  |||  |||  |||  |||  |||
3533 TCATGTTTCACCTTGCTTG...TGGAGATTTCACACTATGC 3569

```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 10, 2002, 00:03:47 ; Search time 1538.44 seconds
(without alignments)
796.273 Million cell updates/sec

Title: US-09-712-768-5
Perfect score: 114
Sequence: 1 atcgctcagcgagccgcaaa.....atgaatcgtccatcacgaa 114

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 11351937 seqs, 537289281 residues
Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST: *
1: em_estfun: *
2: em_esthm: *
3: em_estin: *
4: em_estom: *
5: em_estpl: *
6: em_estba: *
7: em_estlo: *
8: em_estov: *
9: em_hlc: *
10: gb_est1: *
11: gb_est2: *
12: gb_hlc: *
13: gb_gss: *
14: em_gss_fun: *
15: em_gss_hum: *
16: em_gss_inv: *
17: em_gss_pln: *
18: em_gss_pro: *
19: em_gss_rnd: *
20: em_gss_vrt: *
21: em_gss_other: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35.4	31.1	435	11	C82943 C82943 rabb
2	35.4	31.1	435	11	C83799 C83799 rabb
3	34.4	30.2	408	10	BF890606 CM1-MT010
4	34.4	30.2	408	10	AM619231 583 MARC
5	34.4	30.2	454	11	BI185787 UNL-P-FN-
6	34.4	30.2	474	10	AM507072 SOF Band
7	34.4	30.2	482	11	BI337977 361613 MA
8	34.4	30.2	485	10	AM619917 8001 MARC
9	34.4	30.2	486	11	BI337959 361590 MA
10	34.4	30.2	521	11	BI340541 365841 MA
11	34.4	30.2	530	11	BF080960 233775 MA
12	34.4	30.2	539	10	BE231856 BE231856 136697 MA

13	34.4	30.2	590	10	AM619533
14	34.4	30.2	618	10	AM619533
15	34.4	30.2	647	10	AM619624
16	34.4	30.2	781	11	298830
17	33.8	29.6	297	11	BF366213
18	33.8	29.6	837	11	BF43866
19	33.6	29.5	425	10	AM260970
20	33.3	28.9	561	10	AV590980
21	32.8	28.8	397	10	BE032165
22	32.6	28.6	556	10	AM662541
23	32.2	28.2	390	11	BF336164
24	31.4	27.5	503	11	BE18508
25	31.2	27.4	531	11	BF430098
26	31.2	27.4	595	10	AM653359
27	31.1	27.2	302	10	AM414509
28	31.1	27.2	302	10	AM480124
29	31.1	27.2	318	11	BF834501
30	31.1	27.2	354	11	BF834894
31	31.1	27.2	438	11	BF834847
32	31.1	27.2	543	10	AV665622
33	30.6	26.8	366	10	AM480914
34	30.6	26.8	682	11	BF612450
35	30.4	26.7	608	10	AV662352
36	30.4	26.7	644	13	AO957803
37	30.4	26.7	769	13	AO957802
38	30.2	26.5	431	10	AV590979
39	30.2	26.5	453	10	AV589848
40	30.2	26.5	532	10	AV588658
41	30.2	26.5	535	11	BF555571
42	30.2	26.5	550	10	AV591517
43	30.2	26.5	558	10	AV664015
44	30.2	26.5	555	10	AV590621
45	30.2	26.5	564	10	AV589292

ALIGNMENTS

RESULT 1
LOCUS C82943 435 bp mRNA EST 26-MAR-1999
DEFINITION C82943 rabbit corneal endothelial cell Oryctolagus cuniculus CDNA
clone 1414, mRNA sequence.
ACCESSION C82943.1 GI:3061900
VERSION C82943.1
KEYWORDS EST.
SOURCE rabbit.
ORGANISM Oryctolagus cuniculus
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Fukaryota; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
REFERENCE 1 (bases 1 to 435)
AUTHORS Fujimaki, T., Hotta, Y., Sakuma, H., Fujiki, K. and Kanai, A.
TITLE Large-scale sequencing of the rabbit corneal endothelial cDNA library
JOURNAL Cornea 18 (1), 109-114 (1999)
MEDLINE 99110237
COMMENT Contact: Takuro Fujimaki
Department of Ophthalmology
Juntendo University School of Medicine
3-1-3, Hongo, Bunkyo-ku, Tokyo, Japan, Tokyo 113, Japan
Tel: +81-3-5802-1092
Fax: +81-3-5817-0260
Email: fujimaki@med.juntendo.ac.jp.
Location/Qualifiers

FEATURES

source
1..435
/organism="Oryctolagus cuniculus"
/db_xref="taxon:9986"
/clone="1414"
/clone_11b="rabbit corneal endothelial cell"
/cell_type="endothelial cell"
/note="Organ: cornea"
BASE COUNT 109 a 120 c 74 g 132 t
ORIGIN

RESULT	2				
LOCUS	C83799/c				
DEFINITION	C83799	435 bp	mRNA	EST	26-MAR-1999
ACCESSION	C83799	rabbit corneal endothelial cell Oryctolagus cuniculus cDNA			
VERSION	C83799				
KEYWORDS	C83799.1	GI:3062756			
SOURCE	EST.				
	rabbit.				

Department of Ophthalmology
Juntendo University School of Medicine
3-1-3, Hongo, Bunkyo-Ku, Tokyo 113, Japan
Tel.: +81-3-5602-1092
Fax: +81-3-3817-0260
Email: fujimaki@med.juntendo.ac.jp

FEATURES	SOURCE	Location/Qualifiers
	1. . 435	Location/Qualifiers
		/Organism="Oryctolagus cuniculus"
		/db.xref="raxon:3986"
		/clone="1414"
		/clone_1lb="rabbit corneal endothelial cell"
		/cell_type="endothelial cell"
		/note="Organ: cornea"
BASE COUNT	132 a	120 g 109 t
ORIGIN	74 c	

Query Match	31.1%;	Score 35.4;	DB 11;	Length 435;
Best Local Similarity	61.3%;	Pred. No. 0.14;		
Matches	57;	Conservative	0;	Mismatches 26;

	5'	3'	Indels	Gaps
Qy	8	acggcgaccgcaagaataaccgcgatttgcgctagagattgcacgcgcgttgcctgcgatact	67	
Db	381	AAGGCAATCGCAAAAACATACAGCAACGACTCTAGCATTACCAATCTCCCTAGGTATCTTT	322	

QY	68	ttaccctgtgccaa	gctatgat	tattatg	aa	100
Db	321	TTACCC	TACTTCAAG	CATCGAG	ATTACG	AA 289

RESULT	3
BF890606	
LOCUS	BF890606 336 bp
DEFINITION	mrna
ACCESSION	CM1-MT0103-031000-466-h01
VERSION	MT0103
KEYWORDS	Homo sapiens
SOURCE	cdna, mRNA sequence
	BF890606.1 GI:12282065
	EST.
	human.

ORGANISM	REFERENCE	AUTHORS
<i>Homo sapiens</i>		
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.	1 (bases 1 to 336)	Dias Neto, E., Garcia Correa, R., Verlovski-Almeida S. and Santos V. R.

TITLE	JOURNAL	MEDLINE	COMMENT
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)	20202663	Contact: Simpson A.J.G.

Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LIHC Human Cancer Genome
Project. This entry can be seen in the following URL
<http://www.ludwig.org.br/scripts/gethtml2.pl?Cl=CMIc12-CMI-WT0103>
031000-464-h01c1c5=2000-10-03c1c4-1)
Seq primer: puc 18 forward
High quality sequence start: 33
High quality sequence stop: 336.

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="MR0103"
/dev_stage="Adult"
/note="Organ: marrow; Vector: puc18; Site_1: Smal; Site_2: Sma1; A mini-library was made by cloning products derived from ORSITES PCR (U.S. Letters Patent application No. 199716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT
ORIGIN
86 a
125 c
50 g
75 t

```

Query Match	30.28;	Score 34.4;	DB 11;	Length 336;
Best Local Similarity	60.98;	Pred. No. 0.28;		
Matches	56;	Conservative	0;	Mismatches 36;
			Indels	0;
			Gaps	0

Dy 8 acgagcagccgaagaataaacccgcgatctggcctagcatgtcgatcgacgtcgtgatct 67
| | | | | | | | | | | | | | | |
Db 242 AAGSGACCGGAAACATCATTCACAAGCACTATCCATTGCATCACCATTGGACATAGSCGTACT 30

9Y 68 ttaccctctgccaagcctatgaatatattgaa 99
 100 | | | | | | | | | | | | | | | | | |
 Db 302 TCACCCCTCTCCACAGCCTTCAGATATTACGAA 333

RESULT 4
AM619231
DOORS
AMC10002

DEFINITION	408 bp	EST	24-MAR-2000
LOCUS	AW619231	mrna	
DEFINITION	583 MARC PBE Sus scrofa CDNA 5'	mrna sequence.	
ACCSSION	AW619231		
VERSION	AW619231.1	GI:7325415	
KEYWORDS	EST.		
SOURCE	pig.		

ORGANISM
Sus scrofa
Eularyota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
Mammalia: Eutheria: Cetartiodactyla: Suina: Suidae: Sus.
1 (phases 1 to 408)
REFERENCE
AUTHORS
Smith,T.P.L., Fahrenkrug,S.C., Rohrer,G.A., Slimmen,F.A., Rexroad
,C.E. and Keele,J.W.

FEATURES
source
Location/Qualifiers
1. .474

Query Match
Best Local Similarity 30.2%; Score 34.4; DB 10; Length 474;
Matches 56; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="Pig ovarian follicle cDNA"
/sex="female"
/tissue_type="follicle"
/dev_stage="follicular phase"
/note="Organ: ovary; ESTs resulting from differential display experiment"
BASE COUNT 116 a 76 c 145 g 137 t
ORIGIN

RESULT 7
LOCUS B1337977 482 bp mRNA EST 30-JUL-2001
DEFINITION 361613 MARC 1Pig Sus scrofa cDNA 5', mRNA sequence.
ACCESSION B1337977
VERSION B1337977.1 GI:15031260
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 482)
Fahrenkrug, S.C., Freking, B.A., Rohrer, G.A., Smith, T.P.L., Casas, E.,
Stone, R.T., Heaton, M.P., Grose, W.M., Bennett, G.A., Laegreid, W.W.,
and Keefe, J.W.
Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine
Unpublished (2000)
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -mismatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCTCAGTCAGCAG
Plate: 129 row: C column: 18
Seq primer: ATTATGCTGACACTATG.
Location/Qualifiers
1. .482

REFERENCE 1 (bases 1 to 482)
Fahrenkrug, S.C., Freking, B.A., Rohrer, G.A., Smith, T.P.L., Casas, E.,
Stone, R.T., Heaton, M.P., Grose, W.M., Bennett, G.A., Laegreid, W.W.,
and Keefe, J.W.
Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine
Unpublished (2000)
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -mismatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCTCAGTCAGCAG
Plate: 129 row: C column: 18
Seq primer: ATTATGCTGACACTATG.
Location/Qualifiers
1. .482

FEATURES

source
Location/Qualifiers
1. .482

Query Match
Best Local Similarity 30.2%; Score 34.4; DB 10; Length 485;
Matches 56; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 1Pig"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: PCMV SPORT6; Site 1: XbaI; Site 2: XhoI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."
BASE COUNT 119 a 77 c 148 g 138 t
ORIGIN

Query Match
Best Local Similarity 30.2%; Score 34.4; DB 11; Length 482;
Matches 56; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 1Pig"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: PCMV SPORT6; Site 1: XbaI; Site 2: XhoI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."
BASE COUNT 119 a 77 c 148 g 138 t
ORIGIN

Query Match
Best Local Similarity 30.2%; Score 34.4; DB 11; Length 482;
Matches 56; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 1Pig"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: PCMV SPORT6; Site 1: XbaI; Site 2: XhoI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."
BASE COUNT 119 a 77 c 148 g 138 t
ORIGIN

RESULT 8
LOCUS AM619917 485 bp mRNA EST 24-MAR-2000
DEFINITION 8001 MARC PBE Sus scrofa cDNA 5', mRNA sequence.
ACCESSION AM619917
VERSION AM619917.1 GI:7326101
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 485)
Smith, T.P.L., Fahrenkrug, S.C., Rohrer, G.A., Slimen, F.A., Rexroad
Mapping of expressed sequence tags from a porcine early embryonic
cDNA library
Anim. Genet. 32 (2), 66-72 (2001)
21314990
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -mismatch 12 options.
PCR Primers
FORWARD: GGAACAGCTATGACCAT
BACKWARD: GTTTCCTCAGTCAGCAG
Seq primer: ATTATGCTGACACTATG.
Location/Qualifiers
1. .485

FEATURES
source
Location/Qualifiers
1. .485

Query Match
Best Local Similarity 30.2%; Score 34.4; DB 10; Length 485;
Matches 56; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC PBE"
/tissue_type="Day 12 whole embryos"
/lab_host="X10LR"
/note="Vector: pBLUESCRIPT SK-; Site 1: EcoRI; Site 2:
XhoI; Library made from pool of embryos in spherical and
filamentous stages of development (7.58 and 92.5%,
respectively, of each stage) as described in Choi et al,
Endocrinology 137, 1457-67, 1996."
BASE COUNT 146 a 143 c 73 g 123 t
ORIGIN

Query Match
Best Local Similarity 30.2%; Score 34.4; DB 10; Length 485;
Matches 56; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC PBE"
/tissue_type="Day 12 whole embryos"
/lab_host="X10LR"
/note="Vector: pBLUESCRIPT SK-; Site 1: EcoRI; Site 2:
XhoI; Library made from pool of embryos in spherical and
filamentous stages of development (7.58 and 92.5%,
respectively, of each stage) as described in Choi et al,
Endocrinology 137, 1457-67, 1996."
BASE COUNT 146 a 143 c 73 g 123 t
ORIGIN

[illegible]

OM of: US-09-712-768-5 to: SPTREMBL_17.* out_format: pfs
 Date: Jan 10, 2002 1:58 AM

About: Results were produced by the GenCore software, version 4.5,
 Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODL=frnmet_n2p model -DEV=xjh
 -O=/cgn2.1/OSPTO.spool/US09712768/rnmet_09012002_144754_5577/app-query.fasta.1.315
 -DB=SPTREMBL_17 -QFMT=fastan -SUFEX=rspt -GAPOP=12.000
 -GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
 -OGAPEXT=4.500 -OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
 -FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
 -DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosome2
 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
 -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
 -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
 -USER=US09712768 @CGNL_1.103 -NCPU=6 -ICPU=3 -LONGLOG
 -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPRY -WAIT -THREADS=1

Search information block:

Query: US-09-712-768-5
 Query length: 114
 Database: SPTREMBL_17.*
 Database sequences: 473505
 Database length: 146272329
 Search time (sec): 81.490000

score_list:

Sequence	Strid	Orig	ZScore	EScore	Len	Documentation
sp_organelle:Q9MET0	+	84.00	202.11	0.0021	261	Q9MET0 physeter catodon (sperm
sp_organelle:Q9G7T0	+	84.00	201.94	0.0021	267	Q9G7T0 sus scrofa (pig). coli.
sp_organelle:Q9G7S0	+	84.00	201.94	0.0021	267	Q9G7S0 sus scrofa (pig). coli.
sp_organelle:Q9G7S0	+	84.00	201.64	0.0021	278	Q9G7S0 sus scrofa (pig). coli.
sp_organelle:Q9G7S0	+	84.00	200.95	0.0021	305	Q9G7S0 sus scrofa (pig). coli.
sp_organelle:Q9G7S0	+	83.00	199.53	0.0030	261	Q9G7S0 phytophthora infestans (
sp_organelle:Q9G7S0	+	82.00	196.98	0.0041	260	Q9G7S0 rana sylvatica (wood frog
sp_organelle:Q9G7S0	+	82.00	196.95	0.0041	261	Q9G7S0 carnassius auratus (goldf
sp_organelle:Q9G7S0	+	81.00	194.37	0.0057	261	Q9G7S0 salvelinus fontinalis (b
sp_organelle:Q9G7S0	+	81.00	194.37	0.0057	261	Q9G7S0 salvelinus alpinus (arct
sp_organelle:Q9G7S0	+	81.00	194.37	0.0057	261	Q9G7S0 diptophos taenia. cytoch
sp_organelle:Q9G7S0	+	81.00	194.37	0.0057	261	Q9G7S0 echinops telitairi (leese
sp_organelle:Q9G7S0	+	80.00	191.79	0.0080	261	Q9G7S0 isodon macronus (short
sp_organelle:Q9G7S0	+	80.00	191.79	0.0080	261	Q9G7S0 oncorhynchus keta (chum
sp_organelle:Q9G7S0	+	80.00	191.79	0.0080	261	Q9G7S0 oncorhynchus masou (cher
sp_organelle:Q9G7S0	+	80.00	191.79	0.0080	261	Q9G7S0 pleuroglossus altivelis.
sp_organelle:Q9G7S0	+	79.00	189.21	0.0111	261	Q9G7S0 polymyxia japonica. cyto
sp_organelle:Q9G7S0	+	79.00	188.43	0.0110	290	Q9G7S0 caulobacter crescentus.
sp_organelle:Q9G7S0	+	78.00	186.63	0.0154	261	Q9G7S0 cavia porcellus (guinea
sp_organelle:Q9G7S0	+	77.00	190.55	0.0223	109	Q9G7S0 crotaphytus bicornifores.
sp_organelle:Q9G7S0	+	77.00	190.55	0.0223	109	Q9G7S0 crotaphytus collaris. cy
sp_organelle:Q9G7S0	+	77.00	184.05	0.0215	261	Q9G7S0 brachydontio rerio (zebra
sp_organelle:Q9G7S0	+	77.00	184.05	0.0215	261	Q9G7S0 sclurus vulgaris (red sq
sp_organelle:Q9G7S0	+	76.00	184.05	0.0215	261	Q9G7S0 actinoptera tenuis. cytoch
sp_organelle:Q9G7S0	+	75.00	178.89	0.0416	261	Q9G7S0 talco peregrinus. cytoch
sp_organelle:Q9G7S0	+	75.00	178.89	0.0416	261	Q9G7S0 ciconia boyciana. cytoch
sp_organelle:Q9G7S0	+	75.00	178.89	0.0416	261	Q9G7S0 coregonus lavaretus. cyt
sp_organelle:Q9G7S0	+	74.00	176.31	0.0580	261	Q9G7S0 proteolentus dollol (lung
sp_organelle:Q9G7S0	+	74.00	176.31	0.0580	261	Q9G7S0 artibeus jamaicensis (ja
sp_organelle:Q9G7S0	+	74.00	176.31	0.0580	261	Q9G7S0 turnix varis. cytochrome
sp_organelle:Q9G7S0	+	74.00	176.31	0.0580	261	Q9G7S0 eudromia elegans (elegan
sp_organelle:Q9G7S0	+	74.00	176.31	0.0580	261	Q9G7S0 loxodontia africana (afri
sp_organelle:Q9G7S0	+	73.00	179.04	0.0833	128	Q9G7S0 gambelia wislizenii. cyt
sp_organelle:Q9G7S0	+	73.00	173.73	0.0807	261	Q9G7S0 chaetura pelagica (chm
sp_organelle:Q9G7S0	+	73.00	173.73	0.0807	261	Q9G7S0 choroides minor. cytoch
sp_organelle:Q9G7S0	+	73.00	173.73	0.0807	261	Q9G7S0 talpa europaea (european
sp_organelle:Q9G7S0	+	73.00	173.73	0.0807	261	Q9G7S0 satidrops melanostictus.
sp_organelle:Q9G7S0	+	73.00	173.73	0.0807	261	Q9G7S0 autopus japonicus. cytoch
sp_organelle:Q9G7S0	+	72.00	180.34	0.1187	76	Q9G7S0 drosophila adunca (fruit

sp_organelle:Q9G7S0 + 72.00 178.37 0.1173 99 1 Q9G7S0 gambelia wislizenii. c
 sp_organelle:Q9G7S0 + 72.00 178.37 0.1173 99 1 Q9G7S0 gambelia wislizenii. c
 sp_organelle:Q9G7S0 + 72.00 176.46 0.1160 128 1 Q9G7S0 gambelia wislizenii. c
 sp_organelle:Q9G7S0 + 72.00 175.90 0.1156 138 1 Q9G7S0 drosophila phalerata.
 sp_organelle:Q9G7S0 + 72.00 175.90 0.1156 138 1 Q9G7S0 drosophila falleni. c

seq_name: sp_organelle:Q9MET0

seq_documentation_block:

ID Q9MET0 PRELIMINARY; PRT; 261 AA.
 AC Q9MET0;
 DT 01-OCT-2000 (TREMREL. 15, Created)
 DT 01-OCT-2000 (TREMREL. 15, Last sequence update)
 DT 01-JUN-2001 (TREMREL. 17, Last annotation update)
 DE CYTOCHROME OXIDASE SUBUNIT III.
 GN COXIII.
 OS Physeter catodon (Sperm whale) (Physeter macrocephalus).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Odontoceti;
 OC Physeteridae; Physeter.
 OX NCBI_TaxID=9755;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20296818; PubMed=10835487;
 RA Aranson U., Gullberg A., Gretsdotter S., Ursing B., Janke A.;
 RT "The mitochondrial genome of the sperm whale and a new molecular
 reference for estimating eutherian divergence dates.";
 RL J. Mol. Evol. 50:569-578(2000).
 DR EMBL; AJ277029; CAB98276.1; .
 DR InterPro; IPR000298; CytC_oxide_III.
 DR Pfam; PF00510; COX3; 1.
 DR PROSITE; PS0253; COX3; 1.
 KW Mitochondrion.
 SQ SEQUENCE 261 AA: 29907 MW: 87010A5A0AF2F28B CRC64;

alignment_scores:

Quality: 84.00 Length: 37
 Ratio: 3.231 Caps: 0
 Percent Similarity: 70.270 Percent Identity: 51.351

alignment_block:

US-09-712-768-5 x Q9MET0 ..
 Align seg 1/1 to: Q9MET0 from: 1 to: 261

1 ATGTCACGCGCGCAGCAAGAAACCGGATTGGCTTACGATTCGCAT 50
 :::::::::::::::::::::::::::::
 151 LeuMeGlnGlyAspArgLysGlnMetLeuGlnAlaLeuPheIleThrII 167
 :::::::::::::::::::::::::::::::
 51 CGGCGTTGGTGGATCTTACCTGTGCGCAAGCCCTTGAATTTATGAA 100
 :::::::::::::::::::::::::::::::
 167 eataLeuGlyLeuTyPheThrLeuLeuGlnAlaSerGluTyTYrGluII 184
 :::::::::::::::::::::::::::::::
 101 TCGTCCATACC 111
 :::::
 184 hrProPheThr 187

seq_name: sp_organelle:Q9G7T0

seq_documentation_block:

ID Q9G7T0 PRELIMINARY; PRT; 267 AA.
 AC Q9G7T0;
 DT 01-MAR-2001 (TREMREL. 16, Created)
 DT 01-MAR-2001 (TREMREL. 16, Last sequence update)
 DT 01-JUN-2001 (TREMREL. 17, Last annotation update)
 DE COXII.
 GN COXII.
 OS Sus scrofa (pig).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE CYTOCHROME C OXIDASE SUBUNIT 3 (EC 1.9.3.1).
 GN COX3.
 OS Phytophthora infestans (Potato late blight fungus).
 OG Mitochondrion.
 OC Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
 OC Phytophthora.
 OC NCBI_TaxID=4787;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WEST VIRGINIA 4;
 RA Lang B.F., Forget L.;
 RT "The mitochondrial genome of Phytophthora infestans";
 RL Genetic Maps, pp.3:133-135, Cold Spring Harbor Laboratory Press,
 RL NY (1992).
 RL [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=WEST VIRGINIA 4;
 RA Paquin B., Roewer I., Wang Z., Lang B.F.;
 RT "A robust fungal phylogeny using the mitochondrially encoded nad5
 protein sequence";
 RL Can. J. Bot. 73:S180-S185(1995).
 RL [3]
 RN SEQUENCE FROM N.A.
 RC STRAIN=WEST VIRGINIA 4;
 RX MEDLINE=9706266; PubMed=9162109;
 RA Paquin B., Laforest M.J., Forget L., Roewer I., Wang Z., Longcore J.,
 RA Lang B.F.;
 RT The fungal mitochondrial genome project: evolution of fungal
 RT mitochondrial genomes and their gene expression.";
 RL Curr. Genet. 31:380-395(1997).
 RL EMBL; U17009; AAF24786.1;
 DR InterPro: IPR000298; CytC_oxdase_III.
 DR Pfam: PF00510; COX3.1.
 DR PROSITE: PS50253; COX3.1.
 DR Oxidoreductase; Mitochondrion.
 KW SEQUENCE 305 AA; 35262 MW; 72CCE728F7C7CC7 CRC64;

alignment_scores: Length: 33
 Quality: 84.00
 Ratio: 3.111
 Percent Similarity: 81.818 Percent Identity: 54.545

alignment_block:
US-09-712-768-5 x Q9T244 ..

Align seg 1/1 to: Q9T244 from: 1 to: 305

1 ATCTCCACGGCGACGCAAGAAACCGGATTGGCTTACGAGATTGCCAT 50
 ||||| ||||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 194 lIeVAlpHeGlyAspArgLysAsnAlaIleLeuSerLeuIleIleThrl 210
 ||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 51 CGGCTTGCGTGGATCTTTACCTGTCGCAAGCCATGATATATATGAA 99
 ||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 210 eLeuLeuAlaPhePheSerLeuIleGlnAlaIlyrGlyIlyrIleGln 226

seq_name: sp.Organelle:Q9B629

seq_documentation_block:
 ID Q9B629 PRELIMINARY; PRT; 261 AA.
 AC Q9B629;
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE CYTOCHROME C OXIDASE SUBUNIT III.
 GN COIT1
 OS Anguilla japonica (Japanese eel).
 OS Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillidae;

OC Anguillidae; Anguilla.
 ON NCBI_TaxID=7937;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Inoue J.G., Miya M., Nishida M.;
 RT Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RL [2]
 RN SEQUENCE FROM N.A.
 RA Inoue J.G., Miya M., Aoyama J., Ishikawa S., Tsukamoto K., Nishida M.;
 RT "Complete Mitochondrial DNA Sequence of the Japanese Eel, Anguilla
 RT japonica." Fisheries Sci. 67:118-125(2001).
 RL EMBL; AB038556; BAB39718.1;
 DR Mitochondrion.
 KW SEQUENCE 261 AA; 29717 MW; D6317E3C76D4A8A7 CRC64;

alignment_scores: Length: 37
 Quality: 83.00
 Ratio: 2.964
 Percent Similarity: 75.676 Percent Identity: 48.649

alignment_block:
US-09-712-768-5 x Q9B629 ..

Align seg 1/1 to: Q9B629 from: 1 to: 261

1 ATCTCCACGGCGACGCAAGAAACCGGATTGGCTTACGAGATTGCCAT 50
 ||||| ||||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 151 lIeVAlpHeGlyAspArgLysAsnAlaIleGlnSerLeuIleThrl 167
 ||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 51 CGGCTTGCGTGGATCTTTACCTGTCGCAAGCCATGATATATATGAA 100
 ||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 167 eIlleLcnglyPheTyrrPheThrLeuLeuAlaMetIuLyrrGlnA 184
 ||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 101 TCGTCATACC 111
 ||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
 184 lApPheThr 187

seq_name: sp.Organelle:Q9GANU2

seq_documentation_block:
 ID Q9GANU2 PRELIMINARY; PRT; 260 AA.
 AC Q9GANU2;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE CYTOCHROME C OXIDASE POLYPEPTIDE III.
 OS Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Ranidae; Rana.
 ON NCBI_TaxID=45438;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BRAIN;
 RA Wu S.-B., Storey K.B.;
 RT "Up-regulation of mitochondrial ATPase genes under freezing exposure
 RT and anoxic treatment by wood frog (Rana sylvatica).";
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AF175976; AAG33066.1;
 DR InterPro: IPR000298; CytC_oxdase_III.
 DR Pfam: PF00510; COX3.1.
 DR PROSITE: PS50253; COX3.1.
 DR Mitochondrion.
 KW SEQUENCE 260 AA; 29565 MW; 5639DC103F2CD186 CRC64;

alignment_scores: Length: 37
 Quality: 82.00
 Ratio: 3.154
 Percent Similarity: 70.270 Percent Identity: 51.351

alignment_block:

US-09-712-768-5 x Q9GAU2

Align seg 1/1 to: Q9GAU2 from: 1 to: 260

```

1 ATGCTCCAGCGCGCAGCAAGAAACCGGATGCGCTAGCGATGGCCAT 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
151 llemetglnalasparglysglythrleuglnalaleuthrleuthr 167
51 CGGCTTGCTGGATCTTACCTGCGCAAGCCTATGATATATATGAAA 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
167 ethleuglyleutyrrpethrleuenglalameglutrytyrqlua 184
101 TCGTCATACC 111
|||
184 laprophether 187

```

seq_name: sp_organelle:Q9B614

seq_documentation_block:

```

ID Q9B614 PRELIMINARY: PRT: 261 AA.
AC Q9B614;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE CYTOCHROME C OXIDASE SUBUNIT III.
GN CO III.
OS Carassius auratus (Goldfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OX NCBI_TaxID=7957;
RN [1]
RP SEQUENCE FROM N.A.
RA Murakami M.;
RT Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB045144; BAB40354.1; -.
KW Mitochondrion.
SQ SEQUENCE 261 AA; 29658 MW; D0DD83AD5BE184FF CRC64;

```

alignment_scores:

Quality:	Ratio:	Length:
82.00	3.037	37
Percent Similarity: 72.973	Percent Identity: 48.649	Gaps: 0

alignment_block:

US-09-712-768-5 x Q9B614

Align seg 1/1 to: Q9B614 from: 1 to: 261

```

1 ATGCTCCAGCGCGCAGCAAGAAACCGGATGCGCTAGCGATGGCCAT 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
151 llemetglnalasparglysglythrleuenglalameglutrytyrqlua 167
51 CGGCTTGCTGGATCTTACCTGCGCAAGCCTATGATATATATGAAA 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
167 ethleuglyleutyrrpethrleuenglalameglutrytyrqlua 184
101 TCGTCATACC 111
|||
184 laprophether 187

```

seq_name: sp_organelle:Q9ZY39

seq_documentation_block:

```

ID Q9ZY39 PRELIMINARY: PRT: 261 AA.
AC Q9ZY39;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)

```

```

DE CYTOCHROME C OXIDASE SUBUNIT III.
GN COX-III.
OS Rajia radiata.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Rasmussen A.-S.; Arnason U.;
OC Rajiformes; Rajidae; Raja.
OX NCBI_TaxID=27589;
RN [1]
RP SEQUENCE FROM N.A.
RA Rasmussen A.-S.; Arnason U.;
RC MEDLINE=99162577; PubMed=10051614;
RX "Molecular studies suggest that cartilaginous fishes have a terminal
RT position in the piscine tree."
RL Proc. Natl. Acad. Sci. U.S.A. 96:2177-2182(1999).

```

```

RN [2]
RP SEQUENCE FROM N.A.
RA Rasmussen A.-S.;
RC TISSUE=LIVER;
RX Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RL EMBL: AF106038; AAD19935.1; -.
DR HSSP; P18402; 1FFT.
DR InterPro; IPR000298; Cytochrome_c_oxidase_subunit_III.
DR Pfam; PF00510; COX3; 1.
DR PROSITE; PS50253; COX3; 1.
KW Mitochondrion.
SQ SEQUENCE 261 AA; 29600 MW; 6A9B1A6DC779F788 CRC64;

```

alignment_scores:

Quality:	Ratio:	Length:
81.00	3.000	37
Percent Similarity: 72.973	Percent Identity: 45.946	Gaps: 0

alignment_block:

US-09-712-768-5 x Q9ZY39

Align seg 1/1 to: Q9ZY39 from: 1 to: 261

```

1 ATGCTCCAGCGCGCAGCAAGAAACCGGATGCGCTAGCGATGGCCAT 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
151 llemetglnalasparglysglythrleuenglalameglutrytyrqlua 167
51 CGGCTTGCTGGATCTTACCTGCGCAAGCCTATGATATATATGAAA 100
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
167 llemetglnalasparglysglythrleuenglalameglutrytyrqlua 184
101 TCGTCATACC 111
|||
184 laprophether 187

```

seq_name: sp_organelle:Q9XN34

seq_documentation_block:

```

ID Q9XN34 PRELIMINARY: PRT: 261 AA.
AC Q9XN34;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE CYTOCHROME OXIDASE SUBUNIT 3.
GN COIII.
OS Salvelinus fontinalis (Brook trout) (Brook char).
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salvelinus.
OX NCBI_TaxID=8038;
RN [1]
RP SEQUENCE FROM N.A.
RA Doiron S., Blier P.U., Bernatchez L.;
RT "A comparative analysis of complete sequence of mitochondrial genome
between brook char (Salvelinus fontinalis) and arctic char (S.

```

RT alpines).
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF154850; AAD41378.1; .
 DR HSSP: P18402; 1FFT.
 DR InterPro: IPR000298; Cytc_oxdse_III.
 DR Pfam: PF00510; COX3; 1.
 DR PROSITE: PS50253; COX3; 1.
 KW Mitochondrion.
 SQ SEQUENCE 261 AA; 29721 MW; 9EAAAC788FE00316 CRC64;

alignment_scores:
 Quality: 81.00 Length: 37
 Ratio: 2.793 Gaps: 0
 Percent Similarity: 78.378 Percent Identity: 45.946

alignment_block:
 US-09-712-768-5 x Q9XN34 ..

Align seg 1/1 to: Q9XN34 from: 1 to: 261

```

1 ATGCTCACGGCGACCGCAAGAAAACCGCATTTGGCTTACGATTCGCAT 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
151 ILeMetGluGlyGlnArgLysGlnThrValGlnAlaLeuThrLeuThrII 167
51 CGGCCTTGCGCTGGATCTTTACCCCTGTGCGCAAGCCTATGATATATGAAA 100
I |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
167 eLeuLeuGlyPheTyrrPheThrPheLeuGlnGlyMetGluTyrrTyGlnA 184
101 TCGTCCATACC 111
|||
184 laProPheThr 187

```

seq_name: sp.organelle:Q9XN26

seq_documentation_block:
 ID Q9XN26 PRELIMINARY; PRT; 261 AA.
 AC Q9XN26;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE CYTOCHROME OXIDASE SUBUNIT 3.
 GN COXIII.
 OS Salvelinus alpinus (Arctic char).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protactinopterygii; Salmoniformes; Salmonidae; Salvelinus.
 OC NCBITaxID=8036;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Dohren S., Blier P.U., Bernatchez L.;
 RT "A comparative analysis of complete sequence of mitochondrial genome
 RT from brook char (salvelinus fontinalis) and arctic char (S.
 RT alpinus).";
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF154851; AAD41391.1; .
 DR HSSP: P18402; 1FFT.
 DR InterPro: IPR000298; Cytc_oxdse_III.
 DR Pfam: PF00510; COX3; 1.
 DR PROSITE: PS50253; COX3; 1.
 DR Mitochondrion.
 KW Mitochondrion.
 SQ SEQUENCE 261 AA; 29751 MW; A96564CE40C15BA9 CRC64;

alignment_scores:
 Quality: 81.00 Length: 37
 Ratio: 2.793 Gaps: 0
 Percent Similarity: 78.378 Percent Identity: 45.946

alignment_block:
 US-09-712-768-5 x Q9XN26 ..

Align seg 1/1 to: Q9XN26 from: 1 to: 261

```

1 ATGCTCACGGCGACCGCAAGAAAACCGCATTTGGCTTACGATTCGCAT 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
151 ILeMetGluGlyGlnArgLysGlnThrValGlnAlaLeuThrLeuThrII 167
51 CGGCCTTGCGCTGGATCTTTACCCCTGTGCGCAAGCCTATGATATATGAAA 100
I |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
167 eLeuLeuGlyPheTyrrPheThrPheLeuGlnGlyMetGluTyrrTyGlnA 184
101 TCGTCCATACC 111
|||
184 laProPheThr 187

```

seq_name: sp.organelle:Q9G6Q7

seq_documentation_block:
 ID Q9G6Q7 PRELIMINARY; PRT; 261 AA.
 AC Q9G6Q7;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE CYTOCHROME C OXIDASE SUBUNIT III.
 GN COXIII.
 OS Diplophos taenia.
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Stomliiformes; Gonostomatidae; Diplophos.
 OC NCBITaxID=91975;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Miya M., Nishida M.;
 RT "Setting a new stage for the teleostean molecular systematics: A
 RT mtogenomic approach using long PCR and versatile primers.";
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB034825; BAB20726.1; .
 DR InterPro: IPR000298; Cytc_oxdse_III.
 DR Pfam: PF00510; COX3; 1.
 DR PROSITE: PS50253; COX3; 1.
 DR Mitochondrion.
 KW Mitochondrion.
 SQ SEQUENCE 261 AA; 29585 MW; 694856C933298B77 CRC64;

alignment_scores:
 Quality: 81.00 Length: 37
 Ratio: 3.000 Gaps: 0
 Percent Similarity: 72.973 Percent Identity: 48.649

alignment_block:
 US-09-712-768-5 x Q9G6Q7 ..

Align seg 1/1 to: Q9G6Q7 from: 1 to: 261

```

1 ATGCTCACGGCGACCGCAAGAAAACCGCATTTGGCTTACGATTCGCAT 50
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
151 ILeMetAlaGlyGlnArgLysGlnAlaLeuThrLeuThrII 167
51 CGGCCTTGCGCTGGATCTTTACCCCTGTGCGCAAGCCTATGATATATGAAA 100
I |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
167 eLeuLeuGlyPheTyrrPheThrPheLeuGlnAlaLeuLeuTyrrTyGlnA 184
101 TCGTCCATACC 111
|||
184 laProPheThr 187

```

seq_name: sp.organelle:Q9G387

seq_documentation_block:
 ID Q9G387 PRELIMINARY; PRT; 261 AA.
 AC Q9G387;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)

DT 01-JUN-2001 (TREMBlrel. 17, last annotation update)
 DE CYTOCHROME OXIDASE SUBUNIT III.
 GN COIII.
 OS Echinops telfairi (Lesser hedgehog tenrec).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Insectivora; Tenrecidae; Echinops.
 OX NCBI_TaxID=9371;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Mouchaty S.K., Guilberg A., Janke A., Arnason U.;
 RT "Phylogenetic position of the tenrecs (Mammalia: Tenrecidae) of
 RT Madagascar based on analysis of the complete mitochondrial genome
 RT sequence of Echinops telfairi.";
 RL Zool. Scr. 29:307-317(2000).
 DR EMBL: A400734; CAC19398.1; -
 DR InterPro: IPR000298; CYC_c_oxdase_III.
 DR Pfam: PF00510; COX3; 1.
 DR PROSITE: PS50253; COX3; 1.
 KW Mitochondrion.
 SQ SEQUENCE 261 AA; 29773 MW; B66DEB567687DD58 CRC64;

alignment_scores:
 Quality: 81.00 Length: 37
 Ratio: 3.000 Gaps: 0
 Percent Similarity: 72.973 Percent Identity: 48.649

alignment_block:
 US-09-712-768-5 x Q9G387 ..

Align seg 1/1 to: Q9G387 from: 1 to: 261

```

1 ATGCTCAGCGGCGAGCGAAGAAACCGCGATTGGCTACGATGGCCAT 50
  ::::::::::::::::::::::::::::
151 LeuMetGluGlyAsnArgLysGlnMetIleGlnAlaLeuSerIleThrII 167
51 CGGCGCTTGCGTGAGCTTTACCTGTCGCAAGCTATGATATTATGAAA 100
  ::::::::::::::::::::
167 eAlaLeuGlyLeuTyrPheThrIleuGlnAlaSerGluTyrTyGluA 184
101 TCGTCCATACC 111
  |||
184 hApropheThr 187

```

seq_name: sp.organelle:Q9B2G3

seq_documentation_block:
 ID Q9B2G3 PRELIMINARY; PRT: 261 AA.

DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, last sequence update)
 DE CYTOCHROME OXIDASE SUBUNIT III.
 GN COIII.
 OS Isodon macrourus (Short-nosed bandicoot).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Metatheria; Peramelemorphia; Peramelidae; Isodon.
 OX NCBI_TaxID=37698;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Phillips M.J., Lin Y.-H., Harrison G.L., Penny D.;
 RT "Mitochondrial Genomes of a Bandicoot and a Brush-tail Possum Confirm
 RT the Monophyly of Australidelphian Marsupials.";
 RL Proc. R. Soc. Lond., B, Biol. Sci. 0:0-0(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Lin Y.-H., Phillips M.J.;
 RT Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF58864; AK38287.1; -
 KW Mitochondrion.
 SQ SEQUENCE 261 AA; 29966 MW; A108114486943414 CRC64;

alignment_scores:
 Quality: 81.00 Length: 37
 Ratio: 3.000 Gaps: 0
 Percent Similarity: 72.973 Percent Identity: 45.946

alignment_block:
 US-09-712-768-5 x Q9B2G3 ..

Align seg 1/1 to: Q9B2G3 from: 1 to: 261

```

1 ATGCTCAGCGGCGAGCGAAGAAACCGCGATTGGCTACGATGGCCAT 50
  ::::::::::::::::::::::::::::
151 LeuMetGluGlyAsnArgLysGlnMetIleGlnAlaLeuSerIleThrII 167
51 CGGCGCTTGCGTGAGCTTTACCTGTCGCAAGCTATGATATTATGAAA 100
  ::::::::::::::::::::
167 eAlaLeuGlyLeuTyrPheThrIleuGlnAlaMetGluTyrTyGluA 184
101 TCGTCCATACC 111
  |||
184 hApropheThr 187

```

seq_name: sp.organelle:003378

seq_documentation_block:
 ID 003378 PRELIMINARY; PRT: 261 AA.

DT 01-JUL-1997 (TREMBlrel. 04, Created)
 DT 01-JUL-1997 (TREMBlrel. 04, last sequence update)
 DE CYTOCHROME OXIDASE SUBUNIT 3.
 OS Oncorhynchus keta (Chum salmon).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OX NCBI_TaxID=8018;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TOKORO RIVER, HOKKAIDO, JAPAN;
 RX MEDLINE=97159563; PubMed=9007022;
 RA Oonara I., Sawano K., Okazaki T.;
 RT "Mitochondrial DNA sequence analysis of the masu salmon--phylogeny in
 RT the genus Oncorhynchus.";
 RL Mol. Phylogenet. Evol. 7:71-78(1997).
 DR EMBL: D84147; BAA20134.1; -
 DR HSP: P18402; IFT.
 DR InterPro: IPR000298; CYC_c_oxdase_III.
 DR Pfam: PF00510; COX3; 1.
 DR PROSITE: PS50253; COX3; 1.
 KW Mitochondrion.
 SQ SEQUENCE 261 AA; 29691 MW; 4EFES5462ACD66DF CRC64;

alignment_scores:
 Quality: 80.00 Length: 37
 Ratio: 2.857 Gaps: 0
 Percent Similarity: 75.676 Percent Identity: 45.946

alignment_block:
 US-09-712-768-5 x 003378 ..

Align seg 1/1 to: 003378 from: 1 to: 261

```

1 ATGCTCAGCGGCGAGCGAAGAAACCGCGATTGGCTACGATGGCCAT 50
  ::::::::::::::::::::::::::::
151 IleMetGluGlyAsnArgLysGlnMetIleGlnAlaLeuSerIleThrII 167
51 CGGCGCTTGCGTGAGCTTTACCTGTCGCAAGCTATGATATTATGAAA 100
  ::::::::::::::::::::
167 eLeuLeuGlyPheTyrPheThrIleuGlnAlaMetGluTyrTyGluA 184

```


101 TCGTCATACC 111
111
184 laPropherhr 187

Qy 9 cggcgaccgcgaagaacccgagattgagcagatgcagcagccttgatgacatc 68
Db 196 CGGAGACGGGGCACTTCCCATACCGCGGAGGCATCATCGGCTCGGATGGCTTGT 137
Qy 69 taccctgtgcgaacccatgaatatataatgcgtccatc 111
Db 136 GATCTCTCCCTTGTGTCTTCTTCATGAAATCAACATACC 94

RESULT 2

US-08-675-508-21
; Sequence 21, Application US/08675508
; Patent No. 5856136
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; TITLE OF INVENTION: NOVEL HUMAN STEM CELL ANTIGENS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/675,508
; FILING DATE: Filed Herewith
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0066 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 286 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; LIBRARY: UTRSMOT01
; CLONE: 588615
; US-08-675-508-21

Query Match 23.5%; Score 26.8; DB 2; Length 286;
Best Local Similarity 53.6%; Pred. No. 0.85;
Matches 52; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

Qy 13 gaccgcaagaacacgcgattgagcagatgcagcagccttgatgacatc 72
Db 120 GACCGCGGATNCGGAGTTGGCTTCTGACCGTCATCAGCAAGGCTGCGATTGAC 179
Qy 73 ctggtgcaacccatgaatatataatgcgtccatc 109
Db 180 TGCTGTGATGATNACAGGACTACTGCTGCGCAAGA 216

RESULT 3
5187077-16/c
; Patent No. 5187077
; APPLICANT: GEARING, DAVID P.; GOUGH, NICHOLAS M.; HILTON,
; DOUGLAS J.; KING, JULIE A.; METCALF, DONALD; NICE, EDWARD C.;
; NICOLA, NICO A.; SIMPSON, RICHARD J.; WILLSON, TRACY A.
; TITLE OF INVENTION: LEUKEMIA INHIBITORY FACTOR
; NUMBER OF SEQUENCES: 41

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/294,514
; FILING DATE: 09-DEC-1988
; SEQ ID NO: 16
; LENGTH: 1437
5187077-16

Query Match 22.1%; Score 25.2; DB 6; Length 1437;
Best Local Similarity 57.7%; Pred. No. 5.8;
Matches 45; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy 21 gaaacccgattgagcagatgcagcagccttgatgacatc 80
Db 780 GAGTACCTGGAAATCCCTGGGCTGTGCCAGAGCCCGCTTTGCAATTAAACATGTCCCA 721
Qy 81 agcctatgaatatatga 98
Db 720 AACCCAGCAGCATTTTGA 703

RESULT 4

5427925-14/c
; Patent No. 5427925
; APPLICANT: GEARING, DAVID P.; GOUGH, NICHOLAS M.; HILTON,
; DOUGLAS J.; KING, JULIE A.; METCALF, DONALD; NICE, EDWARD C.;
; NICOLA, NICO A.; SIMPSON, RICHARD J.; WILLSON, TRACY A.
; TITLE OF INVENTION: RECOMBINANT METHOD FOR MAKING
; LEUKEMIA INHIBITOR FACTOR
; NUMBER OF SEQUENCES: 38
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/58,979
; FILING DATE: 06-MAY-1993
; APPLICATION NUMBER: 948,614
; FILING DATE: 22-SEP-1992
; APPLICATION NUMBER: 667,159
; FILING DATE: 11-MAR-1991
; SEQ ID NO: 14
; LENGTH: 1437
5427925-14

Query Match 22.1%; Score 25.2; DB 6; Length 1437;
Best Local Similarity 57.7%; Pred. No. 5.8;
Matches 45; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy 21 gaaacccgattgagcagatgcagcagccttgatgacatc 80
Db 780 GAGTACCTGGAAATCCCTGGGCTGTGCCAGAGCCCGCTTTGCAATTAAACATGTCCCA 721
Qy 81 agcctatgaatatatga 98
Db 720 AACCCAGCAGCATTTTGA 703

RESULT 5
US-08-258-261B-6
; Sequence 6, Application US/08258261B
; Patent No. 5639949
; GENERAL INFORMATION:

; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James M.
; APPLICANT: Beck, James Joseph
; APPLICANT: Hill, Dwight Steven
; APPLICANT: Ryals, John Andrew
; APPLICANT: Gaffney, Thomas Deane
; APPLICANT: Lam, Stephen Ting
; APPLICANT: Hammett, Phillip E.
; APPLICANT: Urnes, Scott Joseph
; TITLE OF INVENTION: Genes for the synthesis of
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:

Thu Jan 10 08:15:42 2002

us-09-712-768-5.rn1

Page 3

ADDRESSEE: Ciba-Geigy Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA

ZIP: 10532

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/258,261B

FILING DATE: 08-JUN-1994

CLASSIFICATION: 800

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/457,205

FILING DATE: 01-JUN-1995

ATTORNEY/AGENT INFORMATION:

NAME: Elmer, James Scott

REGISTRATION NUMBER: 36,129

REFERENCE/DOCKET NUMBER: CGC 1506/CIP3

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-541-8614

TELEFAX: 919-541-8689

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 28958 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

US-08-258-261B-6

Query Match 22.1%; Score 25.2; DB 1; Length 28958;
Best Local Similarity 51.8%; Pred. No. 18;
Matches 57; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

OY 3 cgtccagcgacgcaagaacacgcatggtgctgagcgtgctgcatggtgctgctgctg 62

DB 12768 CATCGCGACGCGCCGATGATCTGCCCATGCCGCGCTGCTTCCGCCAAGCCGC 12827

OY 63 gatcttaccctgtgccaagcctatgaatataatgaatgcgtccatccg 112

DB 12828 CAGCGTCCCATCATCTATCTACCGCCTACTATGAGCTGATCTCG 12877

RESULT 6
US-08-456-837-6
Sequence 6, Application US/08456837
Patent No. 5643774

GENERAL INFORMATION:

APPLICANT: Schupp, Thomas

APPLICANT: Ligon, James M.

APPLICANT: Beck, James Joseph

APPLICANT: Hill, Dwight Steven

APPLICANT: Ryals, John Andrew

APPLICANT: Gaffney, Thomas Deane

APPLICANT: Lam, Stephen Ting

APPLICANT: Hammer, Phillip E.

APPLICANT: Uknes, Scott Joseph

TITLE OF INVENTION: Genes for the synthesis of

antipathogenic substances

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: Ciba-Geigy Corporation

STREET: 7 Skyline Drive

CITY: Hawthorne

STATE: NY

COUNTRY: USA

ZIP: 10532

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/456,837

FILING DATE: 01-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/457,205

FILING DATE: 01-JUN-1995

APPLICATION NUMBER: 08/258,261

FILING DATE: 08-JUN-1994

ATTORNEY/AGENT INFORMATION:

NAME: Elmer, James Scott

REGISTRATION NUMBER: 36,129

REFERENCE/DOCKET NUMBER: CGC 1506/CIP3

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-541-8614

TELEFAX: 919-541-8689

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 28958 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

US-08-456-837-6

Query Match 22.1%; Score 25.2; DB 1; Length 28958;
Best Local Similarity 51.8%; Pred. No. 18;
Matches 57; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

OY 3 cgtccagcgacgcaagaacacgcatggtgctgagcgtgctgcatggtgctgctgctg 62

DB 12768 CATCGCGACGCGCCGATGATCTGCCCATGCCGCGCTGCTTCCGCCAAGCCGC 12827

OY 63 gatcttaccctgtgccaagcctatgaatataatgaatgcgtccatccg 112

DB 12828 CAGCGTCCCATCATCTATCTACCGCCTACTATGAGCTGATCTCG 12877

RESULT 7
US-08-457-342-6
Sequence 6, Application US/08457342
Patent No. 5662898

GENERAL INFORMATION:

APPLICANT: Schupp, Thomas

APPLICANT: Ligon, James M.

APPLICANT: Beck, James Joseph

APPLICANT: Hill, Dwight Steven

APPLICANT: Ryals, John Andrew

APPLICANT: Gaffney, Thomas Deane

APPLICANT: Lam, Stephen Ting

APPLICANT: Hammer, Phillip E.

APPLICANT: Uknes, Scott Joseph

TITLE OF INVENTION: Genes for the synthesis of

antipathogenic substances

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: Ciba-Geigy Corporation

STREET: 7 Skyline Drive

CITY: Hawthorne

STATE: NY

COUNTRY: USA

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/457,342
FILING DATE: 01-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/457,205
FILING DATE: 01-JUN-1995
APPLICATION NUMBER: 08/258,261
FILING DATE: 08-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 28958 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-457-342-6

Query Match
Best Local Similarity 22.1%; Score 25.2; DB 1; Length 28958;
Matches 57; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 3 cgtcacgagcgagcgaagaacgcgaltgagcagattgcatcgcccttgctg 62
DB 12768 CATCCGCGAGCCCGCATGATCTGCCCATCCCGCCGCTGCTTCCCAAGCCGC 12827
QY 63 gatcttacccttgccaaagcctatgaatatgaatcgatcgatcaccg 112
DB 12828 CAGCGTCCCATCATCTATCTACCGCCCTACTATGACTGCTCATCTCG 12877

RESULT 8
US-08-457-646A-6
Sequence 6, Application US/08457646A
Patent No. 5679560
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Lyon, James M.
APPLICANT: Beck, James Joseph
APPLICANT: Hill, Dwight Steven
APPLICANT: Ryals, John Andrew
APPLICANT: Gaffney, Thomas Deane
APPLICANT: Lam, Stephen Tling
APPLICANT: Hammer, Phillip E.
TITLE OF INVENTION: Genes for the synthesis of
NUMBER OF INVENTION: antipathogenic substances
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ciba-Geigy Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/457,646A
FILING DATE: 01-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/457,205
FILING DATE: 01-JUN-1995
APPLICATION NUMBER: 08/258,261
FILING DATE: 08-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 28958 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-457-646A-6

Query Match
Best Local Similarity 22.1%; Score 25.2; DB 1; Length 28958;
Matches 57; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 3 cgtcacgagcgagcgaagaacgcgaltgagcagattgcatcgcccttgctg 62
DB 12768 CATCCGCGAGCCCGCATGATCTGCCCATCCCGCCGCTGCTTCCCAAGCCGC 12827
QY 63 gatcttacccttgccaaagcctatgaatatgaatcgatcgatcaccg 112
DB 12828 CAGCGTCCCATCATCTATCTACCGCCCTACTATGACTGCTCATCTCG 12877

RESULT 9
US-08-458-076A-6
Sequence 6, Application US/08458076A
Patent No. 5698425
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Lyon, James M.
APPLICANT: Beck, James Joseph
APPLICANT: Hill, Dwight Steven
APPLICANT: Ryals, John Andrew
APPLICANT: Gaffney, Thomas Deane
APPLICANT: Lam, Stephen Tling
APPLICANT: Hammer, Phillip E.
TITLE OF INVENTION: Genes for the synthesis of
NUMBER OF INVENTION: antipathogenic substances
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ciba-Geigy Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,076A
FILING DATE: 01-JUN-1995

Page 5

Query Match	Score	DB 1	Length
Best Local Similarity	22.1%;	25.2;	28958;
Matches	51.8%;	Pred. No. 18;	
57; Conservative	0;	Mismatches	53;
		Indels	0;
		Gaps	0;

[illegible]

RESULT 11
US-08-457-335A-6
; Sequence 6, Application US/08457335A
; Patent NO. 5723759
; INFORMATION:

US-08-457-335A-6
Sequence 6, Application US/08457335A
Patent No. 5723759
GENERAL INFORMATION:
APPLICANT: Schnpp, Thomas
APPLICANT: Ligon, James M.
APPLICANT: Beck, James Joseph
APPLICANT: Hill, Dwight Steven
APPLICANT: Ryals, John Andrew

GENERAL INFORMATION:
 APPLICANT: Schupp, Thomas
 APPLICANT: Liqon, James M.
 APPLICANT: Beck, James Joseph
 APPLICANT: Hill, Dwight Steven
 APPLICANT: Ryals, John Andrew
 APPLICANT: Gaffney, Thomas Deane
 APPLICANT: Lam, Stephen Ting
 APPLICANT: Hammer, Phillip E.
 APPLICANT: Ukres, Scott Joseph
 TITLE OF INVENTION: Genes for the synthesis of
 TITLE OF INVENTION: antipathogenic substances
 NUMBER OF SEQUENCES: 22
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Ciba-Geigy Corporation

CORRESPONDENCE ADDRESS:
ADDRESS: Ciba-Geigy Corporation
STREET: 7 Skyline Drive
City: Hawthorne
STATE: NY
COUNTRY: USA

```

1 ZIP: 1
2 COMPUTER READABLE FORM:
3 MEDIUM TYPE: Floppy disk
4 COMPUTER: IBM PC compatible
5 OPERATING SYSTEM: PC-DOS/MS-DOS
6 SOFTWARE: patentln Release #1.0, Version #1.25
7 CURRENT APPLICATION DATA:
8 APPLICATION NUMBER: US/08/457,335A
9 FILING DATE: 01-JUN-1995
10 CLASSIFICATION: 800

```

; PRIOR APPLICATION DATE: US 08/457, 205
 ; APPLICATION NUMBER: 01-JUN-1995
 ; FILING DATE: 08/258, 261
 ; APPLICATION NUMBER: 08-Jun-1994
 ; FILING DATE: 08-Jun-1994
 ; ATTORNEY/AGENT INFORMATION:
 ;
 ;
 ;

NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 28958 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-457-335A-6

Query Match
Best Local Similarity 22.1%; Score 25.2; DB 1; Length 28958;
Matches 57; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

Db 12768 CATGCCGACGCCGCGATGTCGCCCATGCCGCGCTGCTCGCCCAAGCCG 12827

QY 63 gatcttaccctgtgccaagcctatgaatattatgaatgcgcataccg 112

Db 12828 CAGCGTCCCATCATCTATCTCACCGCCCTACTATGAGACTGTCGATCTCG 12877

RESULT 12
US-08-729-214-6

Sequence 6, Application US/08729214
Patent No. 5817502

GENERAL INFORMATION:

APPLICANT: Ligon, James M.

APPLICANT: Hill, Dwight Steven

APPLICANT: Ryals, John Andrew

APPLICANT: Hammer, Phillip E.

APPLICANT: Van Pee, Karl-Heinz

APPLICANT: Kliner, Sabine

TITLE OF INVENTION: Genes for the synthesis of

NUMBER OF SEQUENCES: 27

CORRESPONDENCE ADDRESS:

ADDRESSEE: Ciba-Geigy Corporation

STREET: 520 White Plains Road

CITY: Tarrytown

STATE: NY

COUNTRY: USA

ZIP: 10591

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/729,214

FILING DATE: TBA

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Meigs, J. Timothy

REGISTRATION NUMBER: 38,241

REFERENCE/DOCKET NUMBER: CGC1506/CIP5

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-541-8689

TELEFAX: 919-541-8689

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 28958 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-729-214-6

Query Match
Best Local Similarity 22.1%; Score 25.2; DB 1; Length 28958;
Matches 57; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 3 cgtccagcgagccgaagaacgcgagatgctgaagcattgcccagcctgctg 62

Db 12768 CATGCCGACGCCGCGATGTCGCCCATGCCGCGCTGCTCGCCCAAGCCG 12827

QY 63 gatcttaccctgtgccaagcctatgaatattatgaatgcgcataccg 112

Db 12828 CAGCGTCCCATCATCTATCTCACCGCCCTACTATGAGACTGTCGATCTCG 12877

RESULT 13

US-09-028-934-6
Sequence 6, Application US/09028934
Patent No. 6117670

GENERAL INFORMATION:

APPLICANT: Ligon, James M.

APPLICANT: Hill, Dwight S.

APPLICANT: Lam, Steven T.

APPLICANT: Hammer, Phillip E.

APPLICANT: Van Pee, Karl-Heinz

APPLICANT: Kliner, Sabine

APPLICANT: Young, Thomas R.

TITLE OF INVENTION: Pyrolytic Blosynthesis Genes and Uses

NUMBER OF SEQUENCES: 37

CORRESPONDENCE ADDRESS:

ADDRESSEE: No. 6117670artis Corporation

STREET: 3054 Cornwallis Road

CITY: Research Triangle Park

STATE: NC

COUNTRY: USA

ZIP: 27709

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/028,934

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/729,214

FILING DATE: 09-OCT-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/258,261

FILING DATE: 08-JUN-1994

ATTORNEY/AGENT INFORMATION:

NAME: Meigs, J. Timothy

REGISTRATION NUMBER: 38,241

REFERENCE/DOCKET NUMBER: CGC1506/CIP7

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-541-8689

TELEFAX: 919-541-8689

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 28958 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

US-09-028-934-6

Query Match	22.1%;	Score 25.2;	DB 3;	Length 28958;
Best Local Similarity	51.8%;	Pred. NO. 18;		
Matches 57;	Conservative 0;	Mismatches 53;	Indels 0;	Gaps 0

QY 3 cgtccacgagcgaccgaagaacccgcgatcttgctactagcattgccatcgcccttgcctg 62
| | | | | | | | | | | | | | | |
Db 12768 CATCCCGGACGCCCGCATGATCTGGCCCCCATCCCACGCGGTGTCTTCGCGCCAAGCCGC 12827

63 gatctttacccctgccaagccatgatataattargaatcgtccataccg 112
 | | | | | | | | | | | | | | | | | | | | | | | |
 Db 12828 CAGCGTCCCATCATCTATCTCAGCGCCTACTATGAGACTGTCGATCTCG 12877

```

RESULT 14
US-08-764-233A-1
: Sequence 1, Application US/08/64233A
: Patent No. 5716849
:
GENERAL INFORMATION:
: APPLICANT: Ligon, James M.
: APPLICANT: Schupp, Thomas
: APPLICANT: Beck, James J.
: APPLICANT: Hill, Dwight S.
: APPLICANT: Neff, Snezana
: APPLICANT: Ryals, John A.
: TITLE OF INVENTION: Genes For The Biosynthesis Of Soraphen
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Ciba-Geigy Corporation
: STREET: 520 White Plains Road, P.O. Box 2005
: CITY: Tarrytown
: STATE: NY
: COUNTRY: USA
: ZIP: 10591
:
COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/764,233A
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/729,214
: FILING DATE: 09-OCT-1996
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/258,261
: FILING DATE: 08-JUN-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Meigs, J. Timothy
: REGISTRATION NUMBER: 38,241
: REFERENCE/DOCKET NUMBER: 1506/CIP6
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (919) 541-8587
: TELEFAX: (919) 541-8689
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 49377 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: ORIGINAL SOURCE:
: ORGANISM: Sorangium cellulosum
: IMMEDIATE SOURCE:
: CLONE: p98/1, pJL3, and pVKM15
: FEATURE:
: NAME/KEY: misc.feature
: LOCATION: 383..760
: OTHER INFORMATION: /product= "Scop"

```

```

1  /? OTHER INFORMATION: /note= "This gene encodes a protein that is highly homologous to type I PKS such as eryA from Saccharopolyspora erythraea."
2  /? OTHER INFORMATION: the reductase domains of type I PKS
3  /? OTHER INFORMATION: are known to be involved in the synthesis of polyketide compounds."
4  FEATURE:
5  NAME/KEY: misc_feature
6  LOCATION: 927..19874
7  OTHER INFORMATION: /product= "SorA"
8  OTHER INFORMATION: /note= "Gene product is highly homologous to type I PKS
9  OTHER INFORMATION: are known to be involved in the synthesis of polyketide
10 OTHER INFORMATION: compounds."
11 FEATURE:
12 NAME/KEY: misc_feature
13 LOCATION: 942..7115
14 OTHER INFORMATION: /product= "Module 1 of SorA"
15 FEATURE:
16 NAME/KEY: misc_feature
17 LOCATION: 7203..12884
18 OTHER INFORMATION: /product= "Module 2 of SorA"
19 FEATURE:
20 NAME/KEY: misc_feature
21 LOCATION: 13455..19616
22 OTHER INFORMATION: /product= "Module 3 of SorA"
23 FEATURE:
24 NAME/KEY: misc_feature
25 LOCATION: 19871..46318
26 OTHER INFORMATION: /product= "SorB"
27 OTHER INFORMATION: /note= "Gene product is highly homologous to type I PKS
28 FEATURE:
29 NAME/KEY: misc_feature
30 LOCATION: 19870..24556
31 OTHER INFORMATION: /product= "Module 1 of SorB"
32 FEATURE:
33 NAME/KEY: misc_feature
34 LOCATION: 24638..30820
35 OTHER INFORMATION: /product= "Module 2 of SorB"
36 FEATURE:
37 NAME/KEY: misc_feature
38 LOCATION: 30881..35446
39 OTHER INFORMATION: /product= "Module 3 of SorB"
40 FEATURE:
41 NAME/KEY: misc_feature
42 LOCATION: 35528..40114
43 OTHER INFORMATION: /product= "Module 4 of SorB"
44 FEATURE:
45 NAME/KEY: misc_feature
46 LOCATION: 40190..46318
47 OTHER INFORMATION: /product= "Module 5 of SorB"
48 FEATURE:
49 NAME/KEY: misc_feature
50 LOCATION: 46851..47891
51 OTHER INFORMATION: /product= "SorM"
52 OTHER INFORMATION: /note= "The protein encoded by the sorM gene is highly
53 OTHER INFORMATION: homologous to the methyltransferase from Streptomyces
54 OTHER INFORMATION: hygroscopicus that is involved in the synthesis of the
55 OTHER INFORMATION: polyketide rapamycin."
56 US-08-764-233A-1
57 Query Match 22.1%; Score 25.2; DB 1: Length 49377;
58 Best Local Similarity 51.8%; Pred. No. 22;
59 Matches 57; Conservative 0; Mismatches 53; Indels 0; Gaps 0;
60 QY 3 cgtcacggagcgcgaagaacccgcatggatggcgctaggcatgcatacggccttgctg 62
61 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
62 Db 29071 CATGCCGACGCCCGCATGATCTGTCCCATCCCCAGCCCTGGTCCTTCCGCNAAGCCGC 29130
63 gatcttaacctgtgccaaagcctaigaatatgaataatgylcatacgg 112
64 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
65 Db 29131 CAGGTCGCCCATCAWCTATCTCACCGCCTACTATGAGCTGTGATCTMG 29180

```

; Sequence 1, Application US/08920236
; Patent No. 6225083
; GENERAL INFORMATION:
; APPLICANT: Hodgson, John E.
; TITLE OF INVENTION: NOVEL FLSL
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; City: Philadelphia
; STATE: PA
; COUNTRY: US
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTEO for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/920,236
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Dickinson, Todd Q
; REGISTRATION NUMBER: 28,354
; REFERENCE/DOCKET NUMBER: P50556-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2252
; TELEFAX: 215-994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 315 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-920-236-1

Query Match 21.98; Score 25; DB 4; Length 315;
Best Local Similarity 52.4%; Pred. No. 3.9;
Matches 55; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
OY 2 tcgtccacgagcgacgaagaacgcgagttgagcctagcagatgcacatgcgcttgct 61
DB 260 TCTTTCAACGCTTCTGCACGTAATAGTTCATGTGACCTCTTGCTTGCGCATCGTCCAAATTGC 201
OY 62 ggacltaccctgtgccaagcctatgaataatataatgaatcgtcc 106
DB 200 GTTTTCTTTCTCTATCTGCGCATGTGTTTGTCAATCATTC 156

Search completed: January 10, 2002, 01:00:12
Job time: 3180 sec

DR	WPI: 2001-357953/38.
P	P-PSDB: AAY97752.
XX	New cytochrome c oxidase complex having cytochrome c oxidase activity
PT	from Gluconobacter oxydans DSM 4025, useful in mediating electron
PT	transfer in respiratory chain or producing 2-keto-L-gulononic acid from
XX	L-sorbose or D-sorbitol
PS	Claim 15; Page 25; 42pp; English.
XX	This sequence encodes a fragment of the Gluconobacter oxydans
CC	Cytochrome c oxidase, COI1. The invention relates to a cytochrome c
CC	oxidase COI complex. The COI complex is useful in improving oxidative
CC	fermentation and is an essential component mediating electron transfer
CC	in the respiratory chain. The recombinant microorganism and the
CC	cytochrome c oxidase may be used in the genetic preparation of a
CC	recombinant COI complex and in the production of 2-Keto-L-gluconic acid
CC	(2KA) from L-sorbose or D-sorbitol in a culture medium. The COI is also
CC	useful as a terminal oxidase, oxidizing cytochrome c, an electron
CC	acceptor from an enzyme belonging to dehydrogenase for the production of
CC	aldehydes, carboxylic acids and ketones from alcohols and aldehydes,
CC	especially the production of 2KA from L-sorbose or D-sorbitol.
SQ	Sequence 114 BP; 29 A; 33 C; 26 G; 26 T; 0 other;
Query Match	100.0%; Score 114; DB 22; Length 114;
Best Local Similarity	100.0%; Pred. No. 3.4e-32;
Matches 114; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
DY	1 atcgtccacgagaccgcgaagaacacgcgatggcttagcgatgtgcacgtgccttgc 60
Db	1 atcgtccacgagaccgcgaagaacacgcgatggcttagcgatgtgcacgtgccttgc 60
DY	61 tggacttcacctgtgccaaagcctatgaatatgaataatgcttcataccgaa 114
Db	61 tggacttcacctgtgccaaagcctatgaatatgaataatgcttcataccgaa 114
RESULT 2	
AAD00611	
ID	AAD00611 standard; cDNA; 2580 BP.
AC	AAD00611;
DT	29-AUG-2000 (first entry)
DE	Human membrane transport protein, MTRP-12 cDNA.
XX	Human membrane transport protein, MTRP-12 cDNA.
KW	Human; membrane transport protein; MTRP-12; antiinflammatory; cytostatic;
KW	antiHIVid; immunosuppressive; thymometric; antidiabetic; nootropic;
RW	antidiarrhetic; neuroprotective; antidepressant; nephrotoxic; virocid;
KW	antihelmintic; protozoacide; antibacterial; neuroleptic; antigout;
KW	diagnosis; prevention; treatment; membrane transport disorder; epilepsy;
KW	Mekes disease; diabetes; Parkinson's disease; neurological disorder;
KW	Alzheimer's disease; depression; schizophrenia; immune disorder; allergy;
KW	Inflammatory disorder; AIDS; Addison's disease; atherosclerosis; gout;
KW	Graves disease; Hashimoto's thyroiditis; microbial infection; cancer;
XX	cell proliferative disorder; ss.
OS	Homo sapiens.
FH	Key
FT	Location/Qualifiers
FT	CDS 32..2308
FT	/tag= a
XX	/product= "Membrane transport protein-12"
PN	WO200026245-A2.
XX	
XX	11-MAY-2000.
PD	
PF	04-NOV-1999; 99WO-US26048.
XX	
XX	

PR	04-NOV-1998;	9805-0172255.
PR	24-NOV-1998;	9805-0172252.
PR	22-DEC-1998;	9805-0172214.
PR	26-FEB-1999;	9905-0121896.
XX		
PA	(INCY-) INCYTE PHARM INC.	
XX		
PI	Hillman JL, Yue H, Tang YT, Lal P, Corley NC, Guegler KJ,	
PI	Baughn MR, Azimzal Y, Lu DAM;	
XX		
DR	WPI; 2000-365576/31.	
XX	P-PSDB; AAY71067.	
PT		
PT	Novel human membrane transport proteins useful for diagnosis,	
PT	prevention and treatment of membrane transport disorders,	
PT	immune/inflammatory disorders and cell proliferative disorders	
XX	including cancer	
PS		
XX	Claim 9; Page 119-120; 136pp: English.	
XX		
CC	The present sequence is a cDNA encoding membrane transport protein,	
CC	MTRP-12 from Incyte clone 1734477 isolated from human BRSTU8 cDNA	
CC	library. MTRP-12 shows homology to mouse transporter (GI 5359730)	
CC	and human transporter (GI 291964), and is expressed in nervous,	
CC	gastrointestinal and reproductive tissues.	
CC	The present sequence is useful in diagnosis, prevention and treatment	
CC	of disorders related with increased or decreased expression of MTRP	
CC	such as familial galate, Menkes disease, diabetes, Parkinson's disease,	
CC	neurological disorders such as Alzheimer's disease, depression, epilepsy,	
CC	schizophrenia, immune/inflammatory disorders such as AIDS, Addison's	
CC	disease, allergies, atherosclerosis, Graves disease, gout, Hashimoto's	
CC	thyroiditis, viral, bacterial, fungal, parasitic, protozoal or	
CC	hemithic infections and cell proliferative disorders such as cancer.	
CC	Fragments of MTRP polynucleotides are useful as targets in microarrays.	
CC	MTRP DNA is also useful for generating hybridisation probes useful in	
CC	mapping genomic sequences and detecting differences in sequences among	
CC	normal, carrier and affected individuals. It is also used for	
CC	screening libraries of compounds in drug screening techniques.	
XX		
SQ	Sequence 2580 BP; 507 A; 766 C; 743 G; 564 T; 0 other;	
XX		
Query Match	29.6%; Score 33.8; DB 21; Length 2580;	
Best Local Similarity	62.4%; Pred. NO. 0.017;	
Matches 53; Conservative 0; Mismatches 32; Indels 0; Gaps 0;		
QY	2 tcgtccacggcgcgcgcgaagaacccgcagctagcctagcagctgcctgctgc 61	
DB	1062 tcgtgggcagcgcctccacacatcgtctgtggtggttgcagctccatccactggga 1121	
QY	62 ggaacttaccctctgtgccaaagccta 86	
DB	1122 agatcttcgcctcgaagcagcagccta 1146	
RESULT 3		
AAQ070903/c		
ID	AAQ070903 standard; cDNA; 1813 BP.	
XX		
XX	AAQ070903;	
XX		
DT	20-FEB-1995 (first entry)	
XX		
DE	Cold acclimatization protein CAP85 cDNA from spinach leaf.	
XX		
KW	Cold acclimatization; cold-tolerance; transgenic plant; drought-;	
KW	drought-resistance; drought-tolerance; crop improvement; CAP85;	
KW	freezing-tolerance; ss.	
XX		
OS	Spinacia oleracea.	
XX		
XX		
Key	Location/Qualifiers	
CD	34..1641	
TT		

```

FT  XX  /*tag= a
PN  XX  WO9417186-A.
PD  XX  04-AUG-1994.
XX  XX
PF  XX  21-JAN-1994; 94WO-US00581.
XX  XX
PR  XX  21-JAN-1993; 93US-0007107.
XX  XX
PA  XX  (UYFL ) UNITV FLORIDA.
PI  XX  Guy CL, Haskell DW, Hoffig A, Neven LG;
XX  XX  MPI; 1994-264100/32.
DR  XX  P-PSDB: AAR56549.
XX  XX
PT  XX  Cold acclimatization proteins CAP85 and CAP160 and their
PT  XX  nucleotide sequences - used to confer increased cold tolerance
PT  XX  and drought resistance on plants and microorganisms by genetic
PT  XX  transformation
XX  XX
PS  XX  Claim 4; Page 28; 49pp; English.
XX  XX
CC  XX  Transgenic plants constructed using this cold acclimatization
CC  XX  protein nucleotide sequence have improved cold tolerance and/or
CC  XX  drought-resistance. The sequence can also be used as a probe in
CC  XX  assays for crop and plant tolerance levels during seasons of risk to
CC  XX  freezing temp. or drought conditions. The cell transformed with
CC  XX  this sequence is from the family Solanaceae, a citrus plant, a
CC  XX  bacterium or a yeast cell.
XX  XX
SQ  XX  Sequence 1813 BP; 645 A; 392 C; 409 G; 367 T; 0 other:

```

```

Query Match          27.2%; Score 31; DB 15; Length 1813;
Best Local Similarity 56.3%; Pred. No. 0.16;
Matches 58; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

```

```

QY  9  cggcgacgcgaagaacgcgagctgctagcgattgcacatcgcttgctgctatctt 68
    ||| ||| | ||| ||| | ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB  196  GCGAGACGGGCGACACTTCCGATACCGCGAGCGCATCATCGGCTCGATGCTTGT 137
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  69  taccctgtgccaagcctatgaatatatgaatcgtccatcc 111
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB  136  GATCTCTCTCCCTTGTGTTCTTCTTCATGAATCAACATACCC 94
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 4

```

AAK31843
ID  AAK31843 standard; DNA; 1929 BP.
XX
AC  AAK31843;
XX

```

```

DT  14-JUN-1999 (first entry)
XX

```

```

DE  Nitrogen suboxide reductase encoding DNA.
XX

```

```

KW  Nitrogen suboxide reductase; transgenic plant; fortified absorption;
XX  decomposition; nitrogen oxide; ds.
XX

```

```

OS  Achromobacter cycloclastes.
XX

```

```

PN  JP11075842-A.
XX

```

```

PD  23-MAR-1999.
XX

```

```

PF  29-AUG-1997; 97JP-0234200.
XX

```

```

PR  29-AUG-1997; 97JP-0234200.
XX

```

```

PA  (MITO ) MITSUBISHI ELECTRIC CORP.
XX

```

```

DR  WPI; 1999-257693/22.
XX  P-PSDB; AAY03775.
XX  XX
PT  New nitrous oxide reductase gene and protein - useful for production
PT  of transgenic plants with fortified absorption and decomposition of
PT  nitrogen oxide
XX  XX
PS  Claim 2; Pages 7-9; 10pp; Japanese.
XX  XX
CC  This represents a A. cycloclastes nitrogen suboxide reductase encoding
CC  gene. The gene can be used for the production of transgenic plants with
CC  fortified absorption. A microbe and a plant of fortified absorption and
CC  decomposition of nitrogen oxide can be produced by recombining the gene
CC  to the cell of the microbe and the plant.
XX  XX
SQ  Sequence 1929 BP; 427 A; 593 C; 578 G; 331 T; 0 other:

```

```

Query Match          24.4%; Score 27.8; DB 20; Length 1929;
Best Local Similarity 59.5%; Pred. No. 2.4;
Matches 47; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

```

```

QY  7  caccgacgcgaagaacgcgagctgctagcgattgcacatcgcttgctgctatctt 66
    ||| ||| | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB  1288  caaccgacacttgaagagcgtgatggcgaaacgctgatgcgcgaacgactgctc 1347
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY  67  ttaccctgtgccaagcct 85
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB  1348  gttgctgtgcaaatctt 1366
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 5

```

AAV30458
ID  AAV30458 standard; DNA; 534720 BP.
XX

```

```

AC  AAV30458;
XX

```

```

DT  14-OCT-1998 (first entry)
XX

```

```

DE  Rhizobium species plasmid pNGR234a.
XX

```

```

KW  Symbiosis; open reading frame; ORF; plasmid; vector; transportation;
XX  degradation; metabolism; host range; nitrogen fixation; nodulation;
XX

```

```

KM  legume; plant; ds.
XX

```

```

OS  Rhizobium sp.
XX

```

```

XX  Key
XX  Location/Qualifiers
XX

```

```

FH  417796..418671
FT  CDS
XX

```

```

FT  /tag= a
FT  /standard_name= "ORF K1"
FT  /product= "oligopeptide permease"
FT  /note= "homologous to the oppc gene"
XX

```

```

FT  418673..419580
FT  CDS
XX

```

```

FT  /tag= b
FT  /standard_name= "ORF K2"
FT  /product= "oligopeptide permease"
FT  /note= "homologous to the oppd gene"
XX

```

```

FT  419677..420738
FT  CDS
XX

```

```

FT  /tag= c
FT  /standard_name= "ORF K3"
FT  /product= "oligopeptide permease"
FT  /note= "homologous to the oppf gene"
XX

```

```

FT  420774..422159
FT  CDS
XX

```

```

FT  /tag= d
FT  /standard_name= "ORF K4"
FT  /product= "encapsulation-like protein"
FT  /note= "homologous to the CapA gene"
XX

```

```

FT  422628..424031
FT  CDS
XX

```

```

FT  /tag= e
FT  /standard_name= "ORF K5"
FT  /product= "aminotransferase-like protein"
FT  /note= "homologous to the BioA gene"
XX

```

```
FT CDS 424056..425594
FT /tag= f
FT /standard_name= "ORF K6"
FT /product= "(semi)aldehyde dehydrogenase-like protein"
FT CDS complement (426949..428028)
FT /tag= g
FT /standard_name= "ORF K7"
FT /product= "transposase homologue"
FT /note= "homologous to the Trp gene"
FT CDS 428292..429623
FT /tag= h
FT /standard_name= "ORF K8"
FT /product= "glutamate dehydrogenase-like protein"
FT /note= "homologous to the GlutD1 gene"
FT CDS complement (430538..431284)
FT /tag= i
FT /standard_name= "ORF K9"
FT /product= "transposase homologue"
FT CDS complement (431296..432840)
FT /tag= j
FT /standard_name= "ORF K10"
FT /product= "transposase homologue"
FT /note= "homologous to the Trp gene"
FT CDS complement (433880..434110)
FT /tag= k
FT /standard_name= "ORF K11"
FT /product= "protein of unknown function"
FT /note= "homologous to the Fixu gene"
FT CDS complement (434107..434433)
FT /tag= l
FT /standard_name= "ORF K12"
FT /product= "protein of unknown function"
FT CDS complement (434517..434711)
FT /tag= m
FT /standard_name= "ORF K13"
FT /product= "ferrodoxin/ferrodoxin-like protein"
FT /note= "homologous to the FdxN gene"
FT CDS complement (434753..436234)
FT /tag= n
FT /standard_name= "ORF K14"
FT /gene= "nifB"
FT /product= "protein involved in Femo co-factor
FT biosynthesis"
FT CDS complement (436460..438130)
FT /tag= o
FT /standard_name= "ORF K15"
FT /gene= "nifA"
FT /product= "positive regulator of nif, fix and other
FT genes"
FT CDS complement (438297..438590)
FT /tag= p
FT /standard_name= "ORF K16"
FT /gene= "fixX"
FT /product= "protein required for nitrogenase activity"
FT CDS complement (438605..439912)
FT /tag= q
FT /standard_name= "ORF K17"
FT /gene= "fixC"
FT /product= "protein required for nitrogenase activity"
FT CDS complement (439923..441032)
FT /tag= r
FT /standard_name= "ORF K18"
FT /gene= "fixB"
FT /product= "protein required for nitrogenase activity"
FT CDS complement (441042..441899)
FT /tag= s
FT /standard_name= "ORF K19"
FT /gene= "fixA"
FT /product= "protein required for nitrogenase activity"
FT CDS complement (442316..442636)
FT /tag= t
FT /standard_name= "ORF K20"
FT /product= "protein of unknown function"

FT CDS complement (443313..443879)
FT /tag= u
FT /standard_name= "ORF K21"
FT /product= "protein of unknown function"
FT CDS 444337..445029
FT /tag= v
FT /standard_name= "ORF K22"
FT /product= "ferrodoxin-like protein"
FT /note= "homologous to the NifQ gene"
FT CDS 445088..446602
FT /tag= w
FT /standard_name= "ORF K23"
FT /gene= "ncta"
FT /product= "C4-dicarboxylate transport protein"
FT /note= "homologous to the DctAI gene"
FT CDS 446599..447843
FT /tag= x
FT /standard_name= "ORF L1"
FT /product= "cytochrome P450-like protein"
FT /note= "homologous to the CamC gene"
FT CDS 447844..448500
FT /tag= y
FT /standard_name= "ORF L2"
FT /product= "gamma-hexachlorocyclohexane-dechlorinase-like
FT protein"
FT /note= "homologous to the LinA gene"
FT CDS 448497..450203
FT /tag= z
FT /standard_name= "ORF L3"
FT /product= "putative protein with degradative function"
FT CDS 450341..451396
FT /tag= aa
FT /standard_name= "ORF L4"
FT /product= "luciferase alpha-subunit-like protein"
FT /note= "homologous to the LuxA gene"
FT CDS 452980..454494
FT /tag= ab
FT /standard_name= "ORF L6"
FT /gene= "nifD"
FT /product= "alpha-subunit of Femo protein of nitrogenase"
FT 454590..456131
FT /tag= ac
FT /standard_name= "ORF L7"
FT /gene= "nifK"
FT CDS 456187..457677
FT /tag= ad
FT /standard_name= "ORF L8"
FT /product= "protein involved in Femo co-factor
FT biosynthesis"
FT CDS 457687..459096
FT /tag= ae
FT /standard_name= "ORF L9"
FT /product= "protein involved in Femo co-factor
FT biosynthesis"
FT CDS 459093..459575
FT /tag= af
FT /standard_name= "ORF L10"
FT /product= "protein of unknown function"
FT CDS 459579..460067
FT /tag= ag
FT /standard_name= "ORF L11"
FT /product= "protein of unknown function"
FT CDS 460501..460920
FT /tag= ah
FT /standard_name= "ORF L12"
FT /product= "protein similar to part of the Fe protein
FT of nitrogenase"
FT /note= "homologous to the NifH gene"
FT 461228..461545
```

yy

Query Match 23.5%; Score 26.8; DB 19; Length 286;

XX 09-OCT-2001 (first entry)
 XX
 DE Probe #3985 used to measure gene expression in human breast sample.
 XX
 KW Probe: human; breast disease; breast cancer; development disorder; ss;
 XX Inflammatory disease; proliferative breast disease; non-carcinoma tumour.
 OS Homo sapiens.
 XX
 XX WO200157270-A2.
 XX
 XX 09-AUG-2001.
 XX
 XX 29-JAN-2001; 2001WO-US00661.
 XX
 XX 04-FEB-2000; 2000US-0180312.
 XX 26-MAY-2000; 2000US-0207456.
 XX 30-JUN-2000; 2000US-0608408.
 XX 03-AUG-2000; 2000US-0632366.
 XX 21-SEP-2000; 2000US-0234687.
 XX 27-SEP-2000; 2000US-0236359.
 XX 04-OCT-2000; 2000GB-0024263.
 XX
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-476286/51.
 XX
 XX Novel single exon nucleic acid probe used to measuring gene expression
 PT in a human breast -
 XX
 XX Claim 25; SEQ ID No 3985; 322pp; English.
 XX
 XX The present invention relates to novel single exon nucleic acid probes.
 CC The present sequence is one such probe. The probes are useful for
 CC measuring human gene expression in a human breast sample, where the probe
 CC hybridises at high stringency to a nucleic acid expressed in the human
 CC breast. The probes are useful for predicting, diagnosing, grading,
 CC staging, monitoring and prognosing diseases of the human breast,
 CC particularly those diseases with polygenic aetiology. The diseases
 CC include: breast cancer, disorders of development, inflammatory diseases
 CC of the breast, fibrocystic changes, proliferative breast disease and
 CC non-carcinoma tumours.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 CC
 XX Sequence 1959 BP; 521 A; 343 C; 564 G; 531 T; 0 other;

Query Match 23.3%; Score 26.6; DB 22; Length 1959;
 Best Local Similarity 54.6%; Pred. No. 6.5;
 Matches 53; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 1 atgtgcacagcgacgcgaagaaacgcgattgagcctgacgctgagccttggc 60
 DB 443 ATCTATCACCACATCATACCATACCGTCAAGGCTATCATATACCCACATATTGTT 384
 QY 61 tggatcttacctgtgccaagcctatgaatatatg 97
 DB 383 GTCTATCATCATCTCTCAAGCCTATGAGACCTTTG 347

RESULT 11
 ID AAI50726 standard; DNA: 236 BP.
 XX
 XX AAI50726;
 XX
 XX 17-OCT-2001 (first entry)
 XX

DE Probe #19412 used to measure gene expression in human placenta sample.
 XX
 KW Probe: microarray; human; placenta; antenatal diagnosis;
 XX genetic disorder; ss.
 XX
 OS Homo sapiens.
 XX
 XX WO200157272-A2.
 XX
 XX 09-AUG-2001.
 XX
 XX 30-JAN-2001; 2001WO-US00663.
 XX
 XX 04-FEB-2000; 2000US-0180312.
 XX 26-MAY-2000; 2000US-0207456.
 XX 30-JUN-2000; 2000US-0608408.
 XX 03-AUG-2000; 2000US-0632366.
 XX 21-SEP-2000; 2000US-0234687.
 XX 27-SEP-2000; 2000US-0236359.
 XX 04-OCT-2000; 2000GB-0024263.
 XX
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-488897/53.
 XX
 XX Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human placenta -
 XX
 XX Claim 25; SEQ ID No 19412; 654pp; English.
 XX
 XX The present invention relates to single exon nucleic acid probes (SENP).
 CC The present sequence is one such probe. The probes are useful for
 CC producing a microarray for predicting, measuring and displaying gene
 CC expression in samples derived from human placenta. The probes are useful
 CC for antenatal diagnosis of human genetic disorders.
 XX
 XX Sequence 236 BP; 88 A; 60 C; 37 G; 51 T; 0 other;

Query Match 23.0%; Score 26.2; DB 22; Length 236;
 Best Local Similarity 53.4%; Pred. No. 4.4;
 Matches 55; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 6 ccacgagcagccgaagaaacgcgattgagcctgacgcttgcgcttggctgat 65
 DB 79 ccacagcagcgcacaaagaagccttacaagaagccttgatgtatcttcaaccacc 138
 QY 66 cttaacctgtgccaagcctatgaatatatgaatgcgtccat 108
 DB 139 ctgtgactgtgccaactaaataacaaagaagaacacatcat 181

RESULT 12
 ID AAI37551 standard; DNA: 430 BP.
 XX
 XX AAI37551;
 XX
 XX 17-OCT-2001 (first entry)
 XX
 XX Probe #6237 used to measure gene expression in human placenta sample.
 DE Probe: microarray; human; placenta; antenatal diagnosis;
 XX genetic disorder; ss.
 XX
 OS Homo sapiens.
 XX
 XX WO200157272-A2.
 XX
 XX 09-AUG-2001.
 XX

```

PF      30-JAN-2001; 2001WO-US00663.
XX
PR       04-FEB-2000; 2000US-0180312.
XX
PR       26-MAY-2000; 2000US-0207456.
PR       30-JUN-2000; 2000US-0608408.
PR       03-AUG-2000; 2000US-0632366.
PR       21-SEP-2000; 2000US-0234687.
PR       27-SEP-2000; 2000US-0236359.
PR       04-OCT-2000; 2000GB-0024263.
XX
XX      (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI      Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR      WP1; 2001-488897/53.
XX
XX      Human genome-derived single exon nucleic acid probes useful for
PT      analyzing gene expression in human placenta -
XX
PS      Claim 25; SEQ ID No 6237; 654pp; English.
CC
CC      The present invention relates to single exon nucleic acid probes (SNP).
CC      The present sequence is one such probe. The probes are useful for
CC      producing a microarray for predicting, measuring and displaying gene
CC      expression in samples derived from human placenta. The probes are useful
CC      for antenatal diagnosis of human genetic disorders.
XX
SQ      Sequence 430 BP; 154 A; 96 C; 61 G; 119 T; 0 other;

Query Match              23.0%;   Score 26.2; DB 22; Length 430;
Best Local Similarity    53.4%;   Pred. No. 5.4;
Matches 55; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY      6 ccaacgagcaccgcgaagaanaacgcgatggccatgccatggcatggcgctgtgat 65
DB      293 ccacagcacccgcacacaagaaggaccttaaaaaaggccttgattcttccaccacc 352
QY      66 cttagccctggtccaagcctatgaatatataatgaatcgcttcacat 108
DB      353 ctgtgactgtgtgccatcctaataaatatacaaagaacaacactat 395

RESULT 13
AAO50349
ID      AAO50349 standard; DNA; 3129 BP.
XX
AC      AAO50349;
XX
DT      13-MAY-1994 (first entry)
XX
DE      VIP receptor.
XX
KM      vasactive intestinal polypeptide receptor; vip; rat; binding;
adenylyate cyclase activity; stimulus; ds.
OS      Rattus rattus.
XX
FH      Key Location/Qualifiers
FT      CDS           59..1438
          /*tag= a
          /product= VIP_receptor

JP05255394-A.
PN
PD      05-OCT-1993.
XX
PE      13-FEB-1992; 92JP-0026607.
XX
PR      13-FEB-1992; 92JP-0026607.
XX
PA      (OSAB-) ZH OSAKA BIOSCIENCE KENKYUSHO.
```

DR WPI: 1993-348480/44.
 DR P-PSDB: AAR42848.
 XX
 XX Vasoactive intestinal polypeptide - prepd. in large amt. by
 PT culturing microbe transformed by new DNA coding polypeptide
 XX
 XX Claim 4; Page 6; 14pp; Japanese.
 PS
 PS
 CC The sequence can be used to produce large amounts of the VIP
 CC receptor peptide, by culturing a microorganism transformed by
 CC the sequence.
 XX
 XX Sequence 3129 BP; 671 A; 915 C; 852 G; 691 T; 0 other.

Query Match	22.8%	Score 26	DB 14	Length 3129
Best Local Similarity	52.8%	Pred. No. 13		
Matches 56	Conservative 0	Mismatches 50	Indels 0	Gaps 0
QY	6	ccacggcgacgcgaagaacacgcgattgacctgacgatgcatgcgcctgctgcat	65	
Db	265	cctcaccctctgcccacaccccgagggcgagcggtagctctgcatgccccctcat	324	
QY	66	ctttaccctctgccaagcctatgaatataatgaagatgctcatcac	111	
Db	325	ctttcagcgcttgcgcccatctcattggtataataacatcagccgctaac	370	

RESULT	14
AAT96639	
ID	AAT96639 standard; cDNA; 1426 BP.
XX	
AC	AAT96639;
XX	
DT	27-APR-1998 (first entry)
XX	
DE	Human TUB Form 6 cDNA.
XX	
KM	TULP; TUB Form 6; sensory neuron; neurosensory defect;
KW	cochlear degeneration; hearing loss; deafness; retinal dystrophy;
KW	retinitis pigmentosa; combined rod cone dystrophy; obesity; human,
OS	animal model; transgenic animal; therapy; diagnosis; ds.
XX	
XX	Homo sapiens.
FH	
FT	Key Location/Qualifiers
CDS	28..1410
FT	/*tag= a
XX	
PN	WO9738004-A1.
XX	
PD	16-OCT-1997.
XX	
PF	10-APR-1997; 97WO-US05903.
XX	
PR	17-SEP-1996; 96US-0714991.
PR	10-APR-1996; 96US-0630592.
PR	22-AUG-1996; 96US-0701380.
PR	04-SEP-1996; 96US-0706292.
XX	
PPA	(JACK-) JACKSON LAB.
XX	(SEQU-) SEQUANA THERAPEUTICS INC.
PI	Naggert J, Nishina P, Noben-Trauth K, North M;
XX	
DR	WPI; 1997-512642/47.
DR	P-PDB; AAW36488.
XX	
PT	Mammalian TULP protein - used for detecting pre-disposition to
PT	neuro-sensory defects
XX	
VS	Disclosure; Page 48-49; 89pp; English.

Thu Jan 10 08:15:41 2002

us-09-712-768-5.rng

Page 10

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 10, 2002, 00:05:12 ; Search time 1374.03 Seconds
(without alignments)
1368.732 Million cell updates/sec

Title: US-09-712-768-5
Sequence: 1 atcgccacgagcgaccgcaaa.....atgaatcgtccatccgca 114

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GeneBml: *
1: gb_ba: *
2: gb_htg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vi: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_om: *
20: em_or: *
21: em_ov: *
22: em_pat: *
23: em_ph: *
24: em_pl: *
25: em_ro: *
26: em_sts: *
27: em_sy: *
28: em_un: *
29: em_vi: *
30: em_htgo_hum: *
31: em_htgo_inv: *
32: em_htgo_rod: *
33: em_htg_hum: *
34: em_htg_inv: *
35: em_htg_rod: *
36: em_htg_other: *

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	114	100.0	114 6 AX113990	AX113990 Sequence
2	39.8	34.9	346510 1 AP003011	AP003011 Mesorchizo
3	37.6	33.0	17056 4 MTDCNOMGN	Y11832 Dasypus nov
4	37	32.5	16507 10 SVU238588	AJ238588 Sclurus v
5	37	32.5	16968 4 AF348080	AF348080 Ochotona
6	35.6	31.2	783 4 AF030471	AF030471 Gazella s
7	35.6	31.2	1130 4 CA085910	U85910 Ovis aries
8	35.6	31.2	16616 4 AF010406	AF010406 Ovis arie
9	35.4	31.1	2215 4 RABAIAR	L01700 Oryctolagus
10	35.4	31.1	16428 4 PMA277029	AJ277029 Physeter
11	35.4	31.1	15826 4 MIPVDNA	X63726 Phoca vitul
12	35.4	31.1	17245 4 AF030461	AJ001588 Oryctolag
13	35.2	30.5	783 4 AF030472	AF030472 Gazella r
14	34.8	30.5	783 4 AF030473	AF030473 Gazella t
15	34.8	30.5	783 4 AF030470	AF030470 Gazella d
16	34.4	30.2	783 4 AF304202	AF304202 Sus scrof
17	34.4	30.2	15977 4 AF304203	AF304203 Sus scrof
18	34.4	30.2	15977 4 AF304200	AF304200 Sus scrof
19	34.4	30.2	15978 4 AF034253	AF034253 Sus scrof
20	34.4	30.2	16613 4 SS002189	AJ002189 Sus scrof
21	34.4	30.2	17804 5 AF288410	AF288410 Homo sapi
22	34.4	30.0	17804 5 AF288410	AF288410 Homo sapi
23	33.8	29.6	2589 9 AF279265	AF279265 Homo sapi
24	33.8	29.6	2748 9 AC005923	AC005923 Homo sapi
25	33.8	29.6	88326 2 AC024179	AC024179 Homo sapi
26	33.8	29.6	204513 2 AF030464	AF030464 Madagqua g
27	33.2	29.1	783 4 AF030482	AF030482 Gazella d
28	33.2	29.1	783 4 AF030483	AF030483 Gazella d
29	33.2	29.1	783 4 AF030484	AF030484 Gazella d
30	33.2	29.1	783 4 AF030271	AF030271 Trragelaph
31	32.8	28.8	783 4 AF030275	AF030275 Trragelaph
32	32.8	28.8	783 4 AF030278	AF030278 Trragelaph
33	32.8	28.8	783 4 AF030460	AF030460 Ourebla o
34	32.8	28.8	783 4 AF030469	AF030469 Gazella d
35	32.8	28.8	783 4 AF0304201	AF0304201 Sus scrof
36	32.8	28.8	15979 4 AF030457	X70547 Equus cabal
37	32.8	28.8	16660 4 AF030456	AF030456 Aepyceros
38	32.4	28.4	783 4 AF030457	AF030457 Pelica cap
39	32.4	28.4	783 4 AF030474	AF030474 Gazella s
40	32.4	28.4	783 4 AF061340	AF061340 Artibeus
41	32.2	28.2	16651 4 AC073547	AC073547 Homo sapi
42	32.2	28.2	202859 2 AC073547	AC073547 Trragelaph
43	32	28.1	783 4 AF030479	AF030479 Gazella b
44	32	28.1	783 4 AF030481	AF030481 Gazella s
45	32	28.1	783 4 AF030481	AF030481 Gazella s

ALIGNMENTS

RESULT 1
AX113990
LOCUS AX113990 114 bp DNA
DEFINITION Sequence 5 from Patent EP1103603.
ACCESSION AX113990
VERSION AX113990.1 GI:13940150
KEYWORDS
SOURCE
ORGANISM
Glucobacter oxydans.
Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
Glucobacter.
1 (bases 1 to 114)
REFERENCE
Asakura, A., Hoshino, T. and Shinoh, M.
Cytochrome c oxidase complex from glucobacter oxydans
Patent: EP 1103603-A 5 30-MAY-2001;
F. HOFFMANN-LA ROCHE AG (CH)
JOURNAL
FEATURES
source
1. 114
/organism="Glucobacter oxydans"
/db_xref="taxon:442"

```

CDS
1. >114
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAC37871.1"
/db_xref="GI:13940151"
/translation="MVHGRKRTAIGLAIAIGLWIFLCOAYEYEVHTE"

BASE COUNT
29 a 33 c 26 g 26 t

Query Match
Best Local Similarity 100.0%; Score 114; DB 6; Length 114;
Matches 114; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atcgtccagcgagccgaagaacgcgagattggcctaagcattgcacatgcgcctggc 60
Db 1 ATCGTCCAGCGGACCGAGAAACCGCATGGCCATGCGCATGGCCCTTGGC 60

Qy 61 tggattcttacccttgcccaagcctatgatatatgaatgaatcgtccataccga 114
Db 61 TGGATCTTACCCTGTGCCAAGCCTATGATATTTATGAATCGTCATACCGAA 114

RESULT 2
AP003011 346510 bp DNA BCT 15-MAY-2001
DEFINITION Mesorhizobium loti DNA, complete genome, section 18/21.
ACCESSION AP003011 BA000012
VERSION AP003011.2 GI:14026664
KEYWORDS
SOURCE Mesorhizobium loti (strain:MAFF303099) DNA.
ORGANISM Mesorhizobium loti
Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
Phyllobacteriaceae; Mesorhizobium.
1 (sites)
Kaneko, T., Nakamura, Y., Sato, S., Asamizu, E., Kato, T., Sasamoto, S.,
Matanabe, A., Idesawa, K., Ishikawa, A., Kawashima, K., Kimura, T.,
Kishida, Y., Kiyokawa, C., Kohara, M., Matsumoto, M., Matsuno, A.,
Mochizuki, Y., Nakayama, S., Nakazaki, N., Shimpo, S., Sugimoto, K.,
Takeuchi, C., Yamada, M. and Tabata, S.
Complete genome structure of the nitrogen-fixing symbiotic
bacterium Mesorhizobium loti
DNA Res. 7 (6), 331-338 (2000)
2 (bases 1 to 346510)
Kaneko, T.
Direct Submission
Submitted (05-DEC-2000) Takakazu Kaneko, Kazusa DNA Research
Institute, The First Laboratory for Plant Gene Research, Yana
1532-3, Kisarazu, Chiba 292-0812, Japan
(E-mail:kaneko@kazusa.or.jp,
URL:http://www.kazusa.or.jp/rhizobase/,
Tel:81-438-52-3935(ex.2338), Fax:81-438-52-3934)
On May 11, 2001 this sequence version replaced gi:11994986.
Location/Qualifiers
1. 346510
/organism="Mesorhizobium loti"
/strain="MAFF303099"
/db_xref="taxon:381"
/complement(242..1264)
/gene="mlt7087"
/complement(242..1264)
/feature="similar to lipopolysaccharide core biosynthesis
glycosyl transferase Lpse"
/codon_start=1
/transl_table=11
/protein_id="BAB53261.1"
/db_xref="GI:14026665"
/translation="MAGALACAVLVSDVNHPRILAIADRPMAIDRKTONTLRVLER
RFEIVOROHVTSVDVNAADIVITLFWOLAKLPESVLAIRSDLVIGICSHREL
EGEPREROLATLHLRAVAVFNKKLEHEVAPLRLIPVHYTPNGVDITTEFCRPREPQ
CDS
gene
CDS
YSGELRVGASGLGNHGPARRGFHDVIEPANAIPGARILTAIREQMRNHDMLDFY
RDLIVYVCAASRESETPNCTLAACGCIPLVTTRVGNMDELVDGNDLFFEDGAEGLA
DKLALRDTPSLRSMAARKMETLIRGDMWRVQAEVNAHMETSLITLGAAYEPVRRAS
DVRGQT"
complement(1228..2265)
/gene="mlt7088"
/complement(1228..2265)
/gene="mlt7088"
/feature="similar to lipopolysaccharide core biosynthesis
glycosyl transferase Lpse"
/codon_start=1
/transl_table=11
/protein_id="BAB53262.1"
/db_xref="GI:14026666"
/translation="MALINGENHIDLOGGVSTKCGALAMPSPILRLVISAESNCALV
ERPAMHEAVNBEGRNVELNTIHAHAERNSPRVAALTTNRNEERVACLELNFH
OGVOVYCDNOSDRITLQRTYRGAGLIGFESIPNGWYCEQILRRKEELFOSLEA
DMLMHLDDHITLIPYSSYLSVALAADALGFNAIESESEFTPTREAPHONDPO
HTLRTTYPERFSPHCVRAYKKODGPTEIAMSGGHVRVREGPVRLSPRRFKHTFL
SABHAARKYGGRRFLSEEVDRGKWHGMRPLRPEDIQPLDITQVTTATVDLDEANP
WLAHWLARYAS"
complement(2648..3787)
/gene="mlt7089"
/complement(2648..3787)
/gene="mlt7089"
/feature="similar to mandelate racemase"
/codon_start=1
/transl_table=11
/protein_id="BAB53263.1"
/db_xref="GI:14026667"
/translation="MSLSHFRTTRFQEARVDVIGDSQVADNVNAALVELSEGE
IGEITQIENPRLPDQOEITSEVEHEWPSLKNRAIALVHVNPRGNORAYSLPHE
AVOVALMDIAEAKAEGPLHLVLSGRNRNVKAYAGLDFHLDODAPVSLFSAHSTYS
AFKIVGHRDFDRDLRLLELTKTVPAGSKVMIDPNEMATSKALTYIRARGDL
LWVEDPILRHDDGLRLRLHAHVTWTOINSGEYIDLOCKRLLEAHADILNVGQVTD
VMRIQMLAEGLIPISIGNTEFLGAVMAVALPBEVLETSFONFHLVEQPIETIDG
YAYAPDRGHCLVLESEKARGMSRRLANSELGAAPENRILPAK"
3870..4556
/gene="mlt7090"
3870..4556
/gene="mlt7090"
/codon_start=1
/transl_table=11
/product="transcriptional regulator"
/protein_id="BAB53264.1"
/db_xref="GI:14026668"
/translation="MAGKATDSSPLVAARMSDINVERILEGLFDRYPAGAFISON
ELSLIVDPVAPLRDALRVLEAGLITTHPRSGIQVPRGLTETVTOFRSIRRA
VRFPAEGDEAVNNALEIRHSRLRVERDGGKQDLEGMALBEGILGAYIAIRRA
LIDSVYRMRNMYRLRLRLDKITTPIMVRLTKEHLDIEACGTNAADAENKLOHFO
AAMQRSIGLY"
4684..5922
/gene="mlt7091"
4684..5922
/gene="mlt7091"
/codon_start=1
/transl_table=11
/product="sugar ABC transporter, periplasmic binding
protein"
/protein_id="BAB53265.1"
/db_xref="GI:14026669"
/translation="MQLNKQPLMGSAALAMAGHIRGARASATSYWHHTFOSSEMA
LKITIGLOAANADVAVQENIPNSEYKAVSYAVLAGGRDPAATVYAFAPALTYME
GLIDITLDRVNGKRNKAPENRMTGLSNGAIVVPAFADVMWYVYADFEFAGLAG
PRNDEDELISACRKLTDPSKGRVAGEBAGAGGKYVDMEAFSGITVAGQPAIDR
AAAIATITFYSDLFLKEKVVPPSPNSIQRMGCFRTGTAVVWHHTGSLNDSIAM
KOGQESTAPMPAGPKAVAVAVYAGNGIKEDNTEAMRWVSPFQOTDAALTLIEGT
GYFPAASAALDDPTGNPTIHAAVETLLEGRPLNNVGAAGSENVVPAQAVLTG
OSTPEQAVDRITQGLEALR"
5973..6932
/gene="mlt7093"
5973..6932
/gene="mlt7093"
CDS
gene
CDS

```


tRNA
 /gene="tRNA-Val"
 1105..2682
 /product="16S rRNA"
 gene
 1105..2682
 /gene="16S rRNA"
 tRNA
 2684..2758
 /gene="tRNA-Leu (UUR)"
 /product="transfer RNA-Leu"
 gene
 2684..2758
 /product="transfer RNA-Leu"
 gene
 2762..3718
 /gene="tRNA-Leu (UUR)"
 2762..3718
 /gene="NADH1"
 /gene="NADH1"
 2762..3718
 /gene="NADH1"
 /codon_start=1
 /transl_table=2
 /protein_id="CAA72529.1"
 /db_xref="GI:2252501"
 /db_xref="SWISS-PROT:O21325"
 /translation="MYLIVLSLIPILAVAFLLERKILIGWOLKRGPNIVPYG
 LLOPIADAIKLFKEPLRPATSKLMFTLPTLATLALSLMIPMPYPLVNLGV
 LFIAMSLAVYSILMSGNASKYALIGALRAVQITSEVTLAILIILIMTNGSF
 TSLTLETOHMLIFPLPMLMMPITSLAETNAPDPLTEGSESELVSGFNEYAAG
 YPRERDOLMHLKMSFLPTLATLCLMHLVSAFAGVPPHM"
 3717..3785
 /gene="tRNA-Ile"
 /product="transfer RNA-Ile"
 gene
 3717..3785
 /gene="tRNA-Ile"
 complement(3783..3855)
 tRNA
 /product="transfer RNA-Gln"
 complement(3783..3855)
 gene
 3856..3924
 /gene="tRNA-Met"
 /product="transfer RNA-Met"
 gene
 3856..3924
 /gene="tRNA-Met"
 3925..4968
 /gene="NADH2"
 3925..4968
 /gene="NADH2"
 /codon_start=1
 /transl_table=2
 /protein_id="CAA72530.1"
 /db_xref="GI:2252502"
 /db_xref="SWISS-PROT:O21326"
 /translation="NNPLIFIIIMFTLLIGVITMISSHMLIMNGLEMMNESMIPIT
 MMSHPSTEAATKYPMTOATASMLMGVILINLYSGOMTINSGINVTYSMTTAL
 AMKGLAPFHFVPEVTOGLTSGMILITOKLAPMTILQIHSSLMNLMTLAL
 SILIGGNGNOTOLRKIMAYSSIAHMGMTIINTNSLMNLVILMWTITMLAL
 FINSTTTTSLISLTWNSPRTITMTLITLUSGLPLSGFAPAKMIIEMTFNNML
 LPLPTVAMALINLTPYMLITISLTMEPTTNNKMKYKTONFIPLSPTLITST
 MLPLTPMMLILN"
 4967..5034
 /gene="tRNA-Tyr"
 /product="transfer RNA-Tyr"
 gene
 4967..5034
 /gene="tRNA-Tyr"
 complement(5038..5106)
 tRNA
 /gene="tRNA-Ala"
 /product="transfer RNA-Ala"
 complement(5038..5106)
 gene
 /gene="tRNA-Ala"
 complement(5108..5180)
 /product="transfer RNA-Asn"
 complement(5108..5180)
 /gene="tRNA-Asn"
 5185..5218
 rep_origin

tRNA
 /note="origin of L-strand replication"
 complement(5214..5274)
 /gene="tRNA-Cys"
 /product="transfer RNA-Cys"
 gene
 complement(5214..5274)
 /gene="tRNA-Cys"
 tRNA
 complement(5282..5347)
 /gene="tRNA-Tyr"
 /product="transfer RNA-Tyr"
 gene
 complement(5282..5347)
 /gene="tRNA-Tyr"
 5349..6890
 /gene="COI"
 5349..6890
 /gene="COI"
 /codon_start=1
 /transl_table=2
 /protein_id="CAA72525.1"
 /db_xref="GI:2252503"
 /db_xref="SWISS-PROT:O21327"
 /translation="MFTLRLFSTNHRDIDGTYLLFGAWAGWGTALSLITRAELGCP
 GTLLDDQIVNVITVAHAFIMFPMMPIMIGGFWLPLMLGAPDPAFRNNMSP
 WLLPSPSLILASSWVAGAGTGVTPPLACNLAAASVDITITSLDGLGISLIG
 AINFTITIIIMKPPAMTOXOTPLFVSVTVVVLILSLPYLAGITWLLDRNLMT
 FFDPAAGGDPILYQHLFWFGHPREYIILBEGGMI SHIVYSGKKEPEGMGWVA
 MSATIGFGLIYVAHMHMPTVOMDVTNATFTSATMILAIPTGVKFSMLATLHGNIKW
 SPATMLAGTIFRIFTYGGLGMLANSSLDIVLHDTYVVAHFVHVSMAGVFAIMGC
 FVHMFPLFSGTINLTWAKIHFITMVGYNLFFPPOHLGLSGMPRRYSOPPAYTMM
 NYVSSKGSFSLAIVMLIMTMEAFASKREVYVELPTNLMLHCCPPHYHFEED
 AFVKV"
 complement(6892..6960)
 /gene="tRNA-Ser (UCN)"
 /product="transfer RNA-Ser"
 gene
 complement(6892..6960)
 /gene="tRNA-Ser (UCN)"
 6964..7032
 /gene="tRNA-Asp"
 /product="transfer RNA-Asp"
 gene
 6964..7032
 /gene="tRNA-Asp"
 7034..7717
 /gene="COIT"
 7034..7717
 /gene="COIT"
 /codon_start=1
 /transl_table=2
 /protein_id="CAA72526.1"
 /db_xref="GI:2252504"
 /db_xref="SWISS-PROT:P50687"
 /translation="MPPDLGQFODATSPIMEELLFHDHDTMLIFLISLVYITL
 MTKRIHSTMDQAEVETWTLIPAVIILALPSLRILYMMDEINPLDIRAMGH
 QWMSYEITDEDLFDSDYVPTSDIKPGELRLLEVDNVLVPMELSTRMISSEDDY
 HSMVPSLIGKTAIDPRLNOATLMATRGLVGGCSEICSNHSMFIVLELPLAH
 FEMDSTSNL"
 7721..7791
 /gene="tRNA-Lys"
 /product="transfer RNA-Lys"
 gene
 7721..7791
 /gene="tRNA-Lys"
 7792..7995
 /gene="ATPase8"
 7792..7995
 /gene="ATPase8"
 /codon_start=1
 /transl_table=2
 /protein_id="CAA72527.1"
 /db_xref="GI:2252505"
 /db_xref="SWISS-PROT:O21329"
 /translation="MPOLDTSWFTTIVSMLSLFIIMOLKFKIFSSFSFPCPTTMEK
 TKHLTPWEKMKTKTYLPHSLPLP"
 7953..8633
 /gene="ATPase6"
 7953..8633


```

Query Match          33.0%; Score 37.6; DB 4; Length 17056;
Best Local Similarity 63.0%; Pred. No. 0.17;
Matches 58; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

/ gene="ATPase6"
/ codon_start=1
/ trans_table=2
/ protein_id="CAA72528.1"

Oy 8 acggcagccgcaagaacacgcgattggcctagcattgcacgccttgctgcatct 67
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9090 AAGGTACACGGAACACATACCTACTACAGCCCTATTATTACTATCGCTAGCGGTATAC 9149

Oy 68 ttaccctgtccaaagcctatgaatattatgaa 99
    ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9150 TTACCTCTTACAGACCTCAGAAATATTACGAA 9181

RESULT 4
LOCUS SVU238588 16507 bp DNA circular ROD 06-JUN-2000
DEFINITION Sciusurus vulgaris complete mitochondrial genome.
ACCESSION AJ238588
VERSION AJ238588.1 GI:8347068
KEYWORDS 12S ribosomal RNA; 16S ribosomal RNA; 16S rRNA gene;
ATPase 6 gene; ATPase 8 gene; ATPase subunit 6; ATPase subunit 8;
COI gene; COII gene; COIII gene; control region; cytb gene;
cytochrome b; cytochrome c oxidase subunit I; cytochrome c oxidase
subunit II; cytochrome c oxidase subunit III; D-loop;
mitochondrion; NADH dehydrogenase subunit 1; NADH dehydrogenase
subunit 2; NADH dehydrogenase subunit 3; NADH dehydrogenase subunit
4; NADH dehydrogenase subunit 4r; NADH dehydrogenase subunit 5;
NADH dehydrogenase subunit 6; NADH1 gene; NADH2 gene; NADH3 gene;
NADH4 gene; NADH4L gene; NADH5 gene; NADH6 gene; transfer RNA-Ala;
transfer RNA-Arg; transfer RNA-Asn; transfer RNA-Asp; transfer
RNA-Cys; transfer RNA-Gln; transfer RNA-Gly; transfer RNA-His;
transfer RNA-Ile; transfer RNA-Leu(CUN); transfer RNA-Leu(UUR);
transfer RNA-Lys; transfer RNA-Met; transfer RNA-Phe; transfer
RNA-Ser(AGY); transfer RNA-Tyr; transfer RNA-Tyr; transfer RNA-Val;
transfer RNA-Ala gene; transfer RNA-Arg gene; transfer RNA-Asn gene;
transfer RNA-Asp gene; transfer RNA-Cys gene; transfer RNA-Gln gene;
transfer RNA-Gly gene; transfer RNA-His gene; transfer RNA-Ile gene;
transfer RNA-Leu (CUN) gene; transfer RNA-Leu (UUR) gene; transfer
RNA-Lys gene; transfer RNA-Met gene; transfer RNA-Phe gene; transfer
RNA-Ser (AGY) gene; transfer RNA-Tyr gene; transfer RNA-Val gene.

SOURCE Eurasian red squirrel.
ORGANISM Mitochondrion Sciusurus vulgaris
Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Sciuridae; Sciurinae;
Sciurus.
REFERENCE 1 (bases 1 to 16507)
AUTHORS Reyes,A., Gissl,C., Pesole,G., Catzeffis,F.M. and Saccone,C.
TITLE Where do rodents fit? Evidence from the complete mitochondrial
JOURNAL Mol. Biol. Evol. 17 (6), 979-983 (2000)
MEDLINE 20293317
REFERENCE 2 (bases 1 to 16507)
AUTHORS Reyes,A.
TITLE Direct Submission
JOURNAL Submitted (27-Apr-1999) Reyes A., Consiglio Nazionale delle
Ricerche (CNR), Centro Studio Mitochondri e Metabolismo Energetico,
Via Amendola 165/A, Bari, 70126, ITALY

FEATURES
source location/Qualifiers
1..16507
/organism="Sciusurus vulgaris"
/db_xref="mitochondrion"
/db_xref="taxon:55149"
1..66
/ gene="tRNA-Phe"
/ product="transfer RNA-Phe"
1..66
/ gene="tRNA-Phe"
67..1033
/ gene="12S rRNA"

/ product="12S ribosomal RNA"
67..1033
/ gene="12S rRNA"
1034..1102
/ gene="tRNA-Val"
/ product="transfer RNA-Val"
1034..1102
/ gene="tRNA-Val"
1103..2676
/ gene="16S rRNA"
/ product="16S ribosomal RNA"
1103..2676
/ gene="16S rRNA"
2677..2751
/ gene="tRNA-Leu(UUR)"
/ product="transfer RNA-Leu(UUR)"
2677..2751
/ gene="tRNA-Leu(UUR)"
2756..3711
/ gene="NADH1"
/ note="NADH stop codon is completed by the addition of 3' A
residues to the mRNA"
/ codon_start=1
/ trans_table=2
/ product="NADH dehydrogenase subunit 1"
/ protein_id="CAA93985.1"
/ db_xref="GI:8347069"
/ translation="MTLNLFLILIPVIMAFLLVLRKMLGVLQKGNVYPC
LHPPADAKKFLKEPLKSTSVTLFIAPLALTLATITMTIPMPPLNNMGV
LFLATSSLAVALIAMSGLSVALGALRAVQTLSTEVTLAILSVLNNKSF
TSTLTITQCFMWLILPTWPLMMNFIETLAKNAPDLFEGESELVSGNVEYAG
PALFPMAYVTIMIMNALLTFLFGALINPMPPTFTSFALKTLITSTFLMTRAS
YRFKRDQMLHLMKNFLPTLALCMHVSPLIIMACVPPQT"
3712..3780
/ gene="tRNA-Ile"
/ product="transfer RNA-Ile"
3712..3780
/ gene="tRNA-Ile"
/ complement(3778..3849)
/ product="transfer RNA-Gln"
/ complement(3778..3849)
/ gene="tRNA-Gln"
3852..3920
/ gene="tRNA-Met"
/ product="transfer RNA-Met"
3852..3920
/ gene="tRNA-Met"
3921..4962
/ gene="NADH2"
/ note="NADH stop codon is completed by the addition of 3' A
residues to the mRNA"
/ trans_table=2
/ product="NADH dehydrogenase subunit 2"
/ protein_id="CAA93986.1"
/ db_xref="GI:8347070"
/ translation="MNPETSSAIVFTLFGSLITLTFSSRWLIWGLENSLAIIPIL
MSKPNRSIEAASKYFLIQATASIMLMMAIINFTWTGWSISNTINQSLMTPLAL
SMKMGAPFLHVAPEVTOGILPMISGLILTLWOKIAPISIMYQIAPISINYLIMFAMLM
SLILGWGGINOTQRLKILAYSISIAHGMWMAITISNPISLILILYMLINNMEL
LYNKRKTSIALSNSNNKSPILITILMYLMSIGLPLTGFPPKMIIRKEIVSNNI
ILPPLAIVMALMLVIFYRMRLYSTSLTLPSSNNNTIKMLFENTKITPLISALAST
MSLPPLPILSTLY"
4963..5029
/ gene="tRNA-Tyr"
/ product="transfer RNA-Tyr"
4963..5029
/ gene="tRNA-Tyr"

```

Query Match	32.5%;	Score 37;	DB 10;	Length 16507;
Best Local Similarity	62.4%;	Pred. No. 0.26;		
Matches	36;	Conservative	0;	Mismatches 35; Indels 0; Gaps 0
Oy	8	accgacacccgaagaacacgcgatgctgcgcgatccatgcgcctggcctgagatc	67	
Db	9092	AAGGCACCCGACGAAACAAATACCTCCAAAGCAGCTACATTCATCAATTCGACCTTGCCCTCTATT	9151	
Oy	68	ttaccctgtgtccaaacctatgaatatltatgaaa	100	
Db	9152	TTACACTACTCCAAAGCATCCGAATCTACTAGAAA	9184	
RESULT	5			
AF348080				
LOCUS	AF348080	16968 bp	DNA	circular MM
DEFINITION	Ochotona collaris mitochondrion, complete genome.			01-AUG-2001
ACCESSION	AF348080			
VERSION	AF348080.1	GI:14599763		
KEYWORDS				
SOURCE	Ochotona collaris.			
ORGANISM	Mitochondrion Ochotona collaris			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Lagomorpha; Ochotonidae; Ochotona.			
TITLE	Lin, Y.-H., Maddell, P.J. and Penny, D.			
	Pika and vole mitochondrial genomes support both rodent monophyly			
	and girders			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 16968)			
AUTHORS	Lin, Y.-H.			
TITLE	Direct Submission			
JOURNAL	Submitted (08-FEB-2001) Institute of Molecular Biosciences, Massey			
	University, Palmerston North, New Zealand			
FEATURES				
source	Location/Qualifiers			
	1..16968			
	/organism="Ochotona collaris"			
	/organelle="mitochondrion"			
	/db_xref="taxon:134600"			
	1..70			
	/product="tRNA-Phe"			
tRNA	71..1032			
tRNA	/product="12S ribosomal RNA"			
tRNA	1033..1100			
tRNA	/product="tRNA-Val"			
tRNA	1101..2672			
tRNA	/product="16S ribosomal RNA"			
tRNA	2673..2748			
tRNA	/product="tRNA-Leu"			
CDS	2750..3706			
	/codon_start=1			
	/transl_table=2			
	/product="NADH dehydrogenase subunit 1"			
	/protein_id="AAK71069.1"			
	/db_xref="GI:14599764"			
	/translation="MFLAVTILLILIPVLIAAFITLVERRKILTGWOLRKGPNNVGRYG			
	LLOPFAVAALVEKEPRLPSTSPFLFTAPLALTLATLMMIPMPYPLIMNNGI			
	LFLTASLSLAIVSILMSGWSNSKYLALGALRAVAQITSYETLAILLGLLNGSF			
	TLSLSDIDQYKMIILIPAPLMMNFIISLAETNRPADLREGSELSVGFENVEYAG			
	PPALFVLAETNTINMNAITVTLFLGAVHDHNNQLFTINAFITKLTLLTNVFLMITAS			
	YPPRRDQMLHMLKMLKFLPTLITLALCMHVSMPVFLASIPPIA"			
	3706..3773			
tRNA	/product="tRNA-Ile"			
tRNA	complement(3771..3842)			
tRNA	/product="tRNA-Gln"			
tRNA	3843..3911			
tRNA	/product="tRNA-Met"			
CDS	3912..4955			
	/codon_start=1			
	/transl_table=2			
	/product="NADH dehydrogenase subunit 2"			

Best Local Similarity 62.4%; Pred. No. 0.26;
Matches 58; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 8 accgacgcgcgaagaacgcgattgctgagcattgcatcgccctgctgagctc 67
Db 9074 AAGGAACCGCAAAAAACATACACAGACCCCTCTAATATACATCTTACTAGACCCATTT 9133
QY 68 ttaccctgtgccaagcctatgaatatatga 100
Db 9134 TCACCTACTCCAGCCCTCTGATATATGAA 9166

RESULT 6

AF030471 783 bp DNA MAM 26-JUL-1999
LOCUS gazella soemmerringii cytochrome oxidase subunit III (cox III)
DEFINITION gene, mitochondrial gene encoding mitochondrial protein, partial
cds.
AF030471
AF030471.1 GI:2731919
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Soemmerring's gazelle.
Mitochondrion Gazella soemmerringii
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Antilopinae; Gazella.
1 (bases 1 to 783)
Rehholz, W. and Harley, E.
Phylogenetic relationships in the bovid subfamily Antilopinae based
on mitochondrial DNA sequences
Mol. Phylogenet. Evol. 12 (2), 87-94 (1999)
99310773
10381312
2 (bases 1 to 783)
Rehholz, W. E. R. and Harley, E. H.
Direct Submission
Submitted (21-OCT-1997) Chemical Pathology, University of Cape Town
Medical School, Observatory, Cape Town 7925, South Africa
Location/Qualifiers
1..783
/organism="Gazella soemmerringii"
/organism="Mitochondrion"
/sub_species="berberana"
/db_xref="taxon:69306"
/map="8970-9752 relative to the cow (Bos taurus) sequence"
1..>783
/gene="cox III"
1..>783
/gene="cox III"
/codon_start=1
/transl_table=2
/product="cytochrome oxidase subunit III"
/protein_id="AAB93610.1"
/db_xref="GI:2731920"
/translation="MTHTAHYHVNPSWPPLTGALSLMTSGILMFMHNSVALIM
LGLTNNMLTMQWMDVIRESTFGGHHHPNOKGLRYGMILFISEVLFTGPFNAFY
HSLAPRPLGCGWPTGIPHLNPLEVPLNTSVLASGVSITMAHSLMEGRNML
QALFTTALGVYFTLLQSEYEAFPTISDVGSTFPVATGFGHLVITIGSTFLIVC
FEROLKFEHTSSHHFGFEAAAWYHFDVWVLFVLYMWSGCPFSIN"

FEATURES

source
1..783
/organism="Gazella soemmerringii"
/organism="Mitochondrion"
/sub_species="berberana"
/db_xref="taxon:69306"
/map="8970-9752 relative to the cow (Bos taurus) sequence"
1..>783
/gene="cox III"
1..>783
/gene="cox III"
/codon_start=1
/transl_table=2
/product="cytochrome oxidase subunit III"
/protein_id="AAB93610.1"
/db_xref="GI:2731920"
/translation="MTHTAHYHVNPSWPPLTGALSLMTSGILMFMHNSVALIM
LGLTNNMLTMQWMDVIRESTFGGHHHPNOKGLRYGMILFISEVLFTGPFNAFY
HSLAPRPLGCGWPTGIPHLNPLEVPLNTSVLASGVSITMAHSLMEGRNML
QALFTTALGVYFTLLQSEYEAFPTISDVGSTFPVATGFGHLVITIGSTFLIVC
FEROLKFEHTSSHHFGFEAAAWYHFDVWVLFVLYMWSGCPFSIN"

BASE COUNT

213 a 226 c 113 g 231 t

Query Match 31.2%; Score 35.6; DB 4; Length 783;
Best Local Similarity 60.2%; Pred. No. 0.57;
Matches 59; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 2 tcgtcacgcgcgaagaacgcgattgctgagcattgcatcgccctgctgagctc 61
Db 452 TCATAGAGGAGGATCGAACACATCTTCAAGCCCTCTAATATATGACATGACAGAGAG 511
QY 62 ggaattacccgtgccaagcctatgaatatatga 99

Db 512 TGTACTTCACATTAATCTCAAGCCCTCAAGATATATGAA 549

RESULT 7

OA085910 1130 bp DNA MAM 02-OCT-1997
LOCUS Ovis aries cytochrome c-oxidase subunit three (COIII) gene,
DEFINITION complete cds and tRNA-Gly gene, complete sequence, mitochondrial
genes encoding mitochondrial products.
U085910
U085910.1 GI:2459614
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
sheep.
Mitochondrion Ovis aries
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Caprinae; Ovis.
1 (bases 1 to 1130)
Hendleder, S.
Nucleotide sequence of ovine mitochondrial genes for tRNA-Gly and
subunit three of cytochrome c-oxidase
Unpublished
2 (bases 1 to 1130)
Hendleder, S.
Direct Submission
Submitted (21-JAN-1997) Anim. Breed. Genet., JLU-Giessen,
Ludwigstrasse 21B, Giessen D-35390, Germany
Location/Qualifiers
1..1130
/organism="Ovis aries"
/organism="Mitochondrion"
/strain="Merinolandschaf"
/db_xref="taxon:9940"
/tissue_type="liver"
75..878
/gene="COIII"
75..878
/gene="COIII"
/codon_start=1
/transl_table=2
/product="cytochrome c-oxidase subunit three"
/protein_id="AAB71760.1"
/db_xref="GI:2459615"
/translation="MTHTAHYHVNPSWPPLTGALSLMTSGILMFMHNSVALIM
LGLTNNMLTMQWMDVIRESTFGGHHHPNOKGLRYGMILFISEVLFTGPFNAFY
HSLAPRPLGCGWPTGIPHLNPLEVPLNTSVLASGVSITMAHSLMEGRNML
QALFTTALGVYFTLLQSEYEAFPTISDVGSTFPVATGFGHLVITIGSTFLIVC
FEROLKFEHTSSHHFGFEAAAWYHFDVWVLFVLYMWSGCPFSIN"

gene

859..927
/gene="tRNA-gly gene"
859..927
/gene="tRNA-gly gene"
BASE COUNT 307 a 324 c 165 g 334 t

BASE COUNT

307 a 324 c 165 g 334 t

Query Match 31.2%; Score 35.6; DB 4; Length 1130;
Best Local Similarity 60.2%; Pred. No. 0.59;
Matches 59; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 2 tcgtcacgcgcgaagaacgcgattgctgagcattgcatcgccctgctgagctc 61
Db 526 TCATAGAGGAGGATCGAACACATCTTCAAGCCCTCTAATATATGACATGACAGAGAG 585
QY 62 ggaattacccgtgccaagcctatgaatatatga 99
Db 586 TGTACTTCACATTAATCTCAAGCCCTCAAGATATATGAA 623

RESULT 8

AF010406 16616 bp DNA circular MAM 13-DEC-2000

DEFINITION Ovis aries complete mitochondrial genome.
 ACCESSION AF010406
 VERSION AF010406.1 GI:3445513
 KEYWORDS
 SOURCE
 ORGANISM
 sheep.
 Mitochondrion Ovis aries
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Caprinae; Ovis.
 1 (bases 1 to 16616)
 Hiendleder,S., Mainz,K., Planke,Y. and Lewalski,H.
 Analysis of mitochondrial DNA indicates that domestic sheep are
 derived from two different ancestral maternal sources: no evidence
 for contributions from urial and argali sheep
 J. Hered. 89 (2), 113-120 (1998)
 J. Hered. 89 (2), 113-120 (1998)
 2 (bases 1 to 16616)
 Hiendleder,S.
 A low rate of replacement substitutions in two major Ovis aries
 mitochondrial genomes
 Anim. Genet. 23 (2), 116-122 (1998)
 3 (bases 1 to 16616)
 Hiendleder,S., Lewalski,H., Wassmuth,R. and Janke,A.
 The complete mitochondrial DNA sequence of the domestic sheep (Ovis
 aries) and comparison with the other major ovine haplotype
 J. Mol. Evol. 47 (4), 441-448 (1998)
 4 (bases 1 to 16616)
 Hiendleder,S., Wassmuth,R. and Lewalski,H.
 Direct Submission
 Submitted (26-JUN-1997) Animal Breeding and Genetics,
 Justus-Liebig-University, Ludwigstr. 21B, Giessen 35390, Germany
 5 (bases 1 to 16616)
 Hiendleder,S., Wassmuth,R. and Lewalski,H.
 Direct Submission
 Submitted (19-AUG-1998) Animal Breeding and Genetics,
 Justus-Liebig-University, Ludwigstr. 21B, Giessen 35390, Germany
 Sequence update by submitter
 On Aug 21, 1998 this sequence version replaced GI:3366622.
 FEATURES
 source
 1..16616
 /organism="Ovis aries"
 /organelle="mitochondrion"
 /strain="Merinolandschaf"
 /db_xref="taxon:9940"
 /tissue_type="liver"
 /note="haplotype B (European type)"
 1..68
 /product="tRNA-Phe"
 69..1026
 /product="12S ribosomal RNA"
 1027..1093
 /product="tRNA-Val"
 1094..2667
 /product="16S ribosomal RNA"
 2668..2742
 /note="tRNA"
 2745..3699
 /gene="ND1"
 2745..3699
 /gene="ND1"
 /note="ND1"
 /note="tRNA stop codon is completed by the addition of 3' A
 residues to the mRNA"
 /transl_except="(pos:3699,aa:TERM)
 /transl_table=2
 /product="NADH dehydrogenase subunit 1"
 /protein_id="GI:3366623"
 /db_xref="GI:3366623"
 /translation="MFMINVLLIIPILAAVAFLIVLVERKVLGYMGRKRPVAVPYG

LLPPIADAKIKITREKLPATSSISMFILAPILATLALTMWIPMPYPLINMUGV
 LFLMSSLAIVSYILWSCGASNSKVALIGALRAVAOTIEXEVALIILLSVLNGSF
 TLSTLITTOEQVWILTPAMPPLAMMPELSTLAETNPAPDLTEGSEVSGFVVEAAG
 PFLAEFAEYANILIMNIETTLPLGFAHPYPELXTINFTIKSLLSITPLMIRAS
 YPRFYDOLMLMKNFPLPLALCMWHSPLILLSIPPQT
 3701..3769
 /product="tRNA-Ile"
 complement(3767..3838)
 /product="tRNA-Gln"
 3841..3909
 /product="tRNA-Met"
 3910..4951
 /gene="ND2"
 3910..4951
 /gene="ND2"
 /note="tRNA stop codon is completed by the addition of 3' A
 residues to the mRNA"
 /codon_start=1
 /transl_except="(pos:4951,aa:TERM)
 /transl_table=2
 /product="NADH dehydrogenase subunit 2"
 /protein_id="GI:3366624"
 /db_xref="GI:3366624"
 /translation="MAPILITITMTWIGITIVISRWMLLIGFEMNALIPIPM
 MKKNRATVASTKFTFQSTASMLMAIINIMFGQMTVMKLPNMAEMLTAL
 AMKLGAPPHFWPEVTOGIPUSGLITLITOKLAPSVLQIILPSTLIDLITSL
 SITIGMGGLNOTOLKIMAYSSIAHMGMTAVLLYNPTVLLNLIITITMTSGMPL
 FMNSITTTLSLSHTWNAKPIMTIIVLITLISMGILPPLSGMPKMWITIDMTKNDST
 IIPPLMAITALLNLYFYMRITVSTALTMPEPTNNKMKWOPPTTKRMTLTPTWVLSI
 MLPLPILISILE"
 4952..5018
 /product="tRNA-Trp"
 complement(5020..5088)
 /product="tRNA-Ala"
 complement(5090..5162)
 /product="tRNA-Asn"
 5163..5194
 /note="for L-strand"
 complement(5195..5262)
 /product="tRNA-Cys"
 complement(5263..5330)
 /product="tRNA-Tyr"
 5332..6876
 /gene="COI"
 5332..6876
 /gene="COI"
 /codon_start=1
 /transl_table=2
 /product="cytochrome c-oxidase subunit 1"
 /protein_id="GI:3366625"
 /db_xref="GI:3366625"
 /translation="MFIRKRLFSRNHMDIGTLYLLFGAGMAGVSTALSLIRALGCP
 GTLLGDDQIVYIVTAHAFVITFPMKPIIMGCGMNLVPLMIGAPMPEPRNNNSF
 ALNPTFLLILLASSWEGAGTGTVPPLAGNLAHAGASVDLITISILAGVSLIG
 AINFTYLLINNKPPAMSOYQPLFVWSVLTAVLLSLDYLVAAGITMTDGNLMT
 FEDPAGGDPILYQHLPWFPGHPEVYILILFEGMISHIVYISGKEPPTGYGMNTA
 MMSIGELGIYWAHMTFVMDVDVTRAFVSATMIIAIPGVKVFSLATLHGNKTKM
 SPAMMAIGFTELTFTVGGTLGIVLANSIDIVLDITYVAHFIYLSKAVPAIKMG
 FVHMFPLPSGTYLNTDTPAKIHFAIMFVGNMTFFPQHLGSGMPRRYSYDPAYTMM
 NTISSGFSISLAVMLMTIITWEAFASKREVLYVDLTTLNLEWLNCCPPHYHTFEP
 TYVNLK"
 complement(6874..6944)
 /note="UCN"
 /product="tRNA-Ser"
 6950..7017
 /product="tRNA-Asp"
 7019..7702
 /gene="COI1"
 7019..7702
 /gene="COI1"
 /codon_start=1
 /transl_table=2
 /product="cytochrome c-oxidase subunit 2"

/protein_id="AADI0099.1"
 /db_xref="GI:336626"
 /translation="MAYPMQGFODATSPIMPELLHEDHMLIVELLISLVYIISL
 MUTTKLHTSTDAOEVEITITLPAIILIMALESLITLMDETNPSLFTVGMH
 OMVASYETDYEDSPDSYMITPSELKGELEKLEVDNRVVLPMEMVPMIISSEDL
 PSWAVPSGLKTDALPGRNLQTTMLSTRKGLFYGQCEISCHNSHFMPIVLEVLAKY
 FKWASASML"
 7706..7773
 /product="tRNA-Lys"
 /gene="tRNA-Lys"
 7775..7975
 /gene="Atpase8"
 7775..7975
 /gene="Atpase8"
 /codon_start=1
 /transl_table=2
 /product="Atpase subunit 8"
 /protein_id="AADI0100.1"
 /db_xref="GI:336627"
 /translation="MPQDSTWMLTMSLMLVLIIFQIKSKHNFHNPDLMTYKT
 PKQMPWETKTKTYLPLSLPL"
 7936..8614
 /gene="Atpase6"
 7936..8614
 /gene="Atpase6"
 /note="TAA stop codon is completed by the addition of 3' A
 residues to the mRNA"
 /codon_start=1
 /transl_except=(pos:8614,aa:TERM)
 /transl_table=2
 /product="Atpase subunit 6"
 /protein_id="AADI0101.1"
 /db_xref="GI:336628"
 /translation="MNENLPSFIPPMGCLPLVLYLFPSELPTSNRLVNNRLIS
 LQQMLQVSKOMSTHNTKOTWALMLSLILFISTNLGLPLHSTPTQSLML
 GMAIPMGAVITGFNRKTRKASLHFLPGSTPLIPMLVLIETISLFIQVVALAVRL
 SYLHNDNT"
 8616..9399
 /gene="COI1I"
 8616..9399
 /gene="COI1I"
 /note="TAA stop codon is completed by the addition of 3' A
 residues to the mRNA"
 /codon_start=1
 /transl_except=(pos:9399,aa:TERM)
 /transl_table=2
 /product="cytochrome c-oxidase subunit 3"
 /protein_id="AADI0108.1"
 /db_xref="GI:336635"
 /translation="MTHTHAYHVNPSWPLGALSALIMTSGLIWMFHNSTALLT
 LGITNNMLITMOWKSDVIRESTFGHHHTPAVOKLARGMTLITSEVLPTGPMFV
 HSSLPVTELGCPMPCGIIHPLNLEPLPLNTSVLASGSIIMAHHSIMEGRYML
 OALFTTALGVFTLLQASEYEFAPFTISDGVGSTFVFATGHHGLVITGTFLLVC
 31.2%; Score 35.6; DB 4; Length 16616;
 Best Local Similarity 60.2%; Pred. No. 0.74;
 Matches 59; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
 QY 2 tcgtcacggcgaccgcaagaacacggcattgctgctagcattgcacgctgctgct 61
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 9067 TCATGAAGGAGGACCTTACACATGTTCACAAAGCCCTATTCATCAGCATGCACACTAGCGC 9126
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 62 ggaatcttaccctgtgccaagcctatgaattatga 99
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 9127 TGTACTTACACTGTTCACAGCATCAGATATATGAA 9164
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 RESULT 9
 LOCUS RABAIAR 2215 bp mRNA MAM 04-AUG-1993
 DEFINITION Oryctolagus cuniculus A1 adenosine receptor mRNA, complete cds.
 ACCESSION L01700
 VERSION L01700.1 GI:347705

KEYWORDS A1 adenosine receptor.
 SOURCE Oryctolagus cuniculus (library: lambda ZAP11/EMBL3) cDNA to mRNA.
 ORGANISM Oryctolagus cuniculus
 BUKARYOTA; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 REFERENCE 1 (bases 1 to 2215)
 Bhattacharya,S., Dewitt,D.L., Burnatowska-Hledin,M., Smith,W.L. and
 Spielman,W.S.
 Cloning of an adenosine A-1 receptor-encoding gene from rabbit
 Gene 128, 285-286 (1993)
 FEATURES
 source
 1..2215
 /organism="Oryctolagus cuniculus"
 /db_xref="taxon:9986"
 /tissue="lib="lambda ZAP11/EMBL3"
 1..378
 /number=1
 38..1024
 /codon_start=1
 /product="A1 adenosine receptor"
 /protein_id="AA31148.1"
 /db_xref="GI:347706"
 /translation="MPPSISAFQAGVIGIEVLATVSPGNVLTMAVKNQALBDAT
 FCRVSLADVADVAGLVLPALILINIGETVPHICIMVACPVILIPQSLILALTA
 VDRILRKIPRLRYKAVTPRAVAIAGCMISLVGLTPMGKNNLENRANRANG
 SVSEPVYKCEFEKVIISMEYVYFVFWVLPPLMLIYLEVPYLIRQLSKRASAS
 SGDPHRYKGEKELIAKSLALILFLPLSMPLHITNCVLTLCPSQKPSILVYAIPL
 THNSAMNPVIVAFRIHFRVRYEIKIWNDFHFCRPAAGDDEDLPEKPNQ"
 BASE COUNT 401 a 720 c 644 g 450 t
 ORIGIN
 Query Match 31.1%; Score 35.4; DB 4; Length 2215;
 Best Local Similarity 61.3%; Pred. No. 0.72;
 Matches 57; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
 QY 8 acggcgaccgcaagaacacggcattgctgctagcattgcacgctgctgctgct 67
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1892 AAGGCATCGCAAAACATACACGACGCTGTAGCCATTCCTCCTAGGTATCATTT 1951
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 68 ttaccctgtgccaagcctatgaattatga 100
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1952 TTACCTTACTTCAAGCATCGAGTATTACGANA 1984
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 RESULT 10
 LOCUS PMA277029 16428 bp DNA circular MAM 20-JUL-2000
 DEFINITION Physeter macrocephalus mitochondrial genome.
 ACCESSION AJ277029
 VERSION AJ277029.1 GI:9368962
 KEYWORDS 12S ribosomal RNA; 16S rRNA gene; 16S rRNA gene;
 Atpase 6 gene; Atpase 8 gene; Atpase subunit 6; Atpase subunit 8;
 COI gene; COII gene; COIII gene; cytochrome b; cytochrome b gene;
 cytochrome oxidase subunit I; cytochrome oxidase subunit II;
 dehydrogenase 1; NADH dehydrogenase subunit 3; NADH dehydrogenase
 subunit 4; NADH dehydrogenase subunit 6; NADH dehydrogenase
 subunit 5; NADH dehydrogenase subunit 4L; NADH dehydrogenase
 2; NADH1 gene; NADH2 gene; NADH3 gene; NADH4 gene; NADH4L gene;
 NADH5 gene; NADH6 gene; transfer RNA-Ala; transfer RNA-Asn;
 transfer RNA-Asp; transfer RNA-Cys; transfer RNA-Gln; transfer
 RNA-Glu; transfer RNA-Gly; transfer RNA-His; transfer RNA-Ile;
 transfer RNA-Leu(CUN); transfer RNA-Leu(UUR); transfer RNA-Lys;
 transfer RNA-Met; transfer RNA-Phe; transfer RNA-Pro; transfer
 RNA-Ser(AGY); transfer RNA-Ser(UCN); transfer RNA-Thr; transfer
 RNA-Trp; transfer RNA-Tyr; transfer RNA-Val.
 SOURCE
 ORGANISM Mitochondrion Physeter catodon
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Cetacea; Odontoceti;
 Physeteridae; Physeter.

[illegible][illegible]


```

/gene="tRNA-Asp"
/product="transfer RNA-Asp"
6970..7038
/gene="tRNA-Asp"
7039..7722
/gene="ccolI"
7039..7722
/gene="ccolII"
/codon_start=1
/transl_table=2
/product="cytochrome c oxidase subunit II"
/protein_id="CA04850.1"
/db_xref="GI:3293010"
/db_xref="SPTREMBL:O79430"
/translation="MAVFFQGFODASPIHEELLHFHDHMLIVFLISLVYISL
MTRKLTPTSTMDAOEVEETITWILPAIILIALPSLIYMDEINNPISLVTMGH
OMTMYEITYEDLANFDSYMTPTSDINGDRLIEVDNRVLPMPELIRMISSDYLA
HSMAVPSGLKTDAIPGRLNQATILSTRPGLFYQCSCSLCSNHSFPIYLEAVPLKH
FENMSLSMI"
tRNA
7726..7794
/product="transfer RNA-Lys"
7726..7794
gene
Query Match      31.1%; Score 35.4; DB 4; Length 17245;
Best local Similarity 61.3%; Pzed. No. 0.86; Mismatches 36; Indels 0; Gaps 0;
Matches 57; Conservative
Oy 8 acggcagccgcaagaanaacgcgatgtgcgcattgccatcgccatgcgctgctgcatct 67
    | ||| | ||||| | ||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Db 9094 AAGGCATGCCAANAACATTCACGACMACCTTACCACATTCCATTCCTAGTACTATT 9153
    | ||| | ||||| | ||| | ||||| | ||||| | ||||| | ||||| | ||||| |
Oy 68 ttaccctgtccaagccataatgaatattgaaa 100
    ||||||| | ||||| | || ||||| |||||
Db 9154 TTACCCTACTTCAAGCATCGAGATTATCGAAA 9186
    ||||||| | ||||| | || ||||| |||||

RESULT 13
AF030461          783 bp      DNA       NAM       26-JUL-1999
LOCUS             Raphicerus campestris cytochrome oxidase subunit III (cox iii)
DEFINITION        gene, mitochondrial gene encoding mitochondrial protein, partial
                   c88
ACCESSION         AF030461
VERSION           AF030461.1   GI:2731899
KEYWORDS
SOURCE            steenbok.
ORGNAMISM         Mitochondrion Raphicerus campestris
                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
                  Bovidae; Antilopinae; Raphicerus.
REFERENCE          1 (bases 1 to 783)
AUTHORS           Rebholz,W. and Harley,E.
TITLE              Phylogenetic relationships in the bovid subfamily Antilopinae based
                   on mitochondrial DNA sequences
JOURNAL            Mol. Phylogenet. Evol. 12 (2), 87-94 (1999)
MEDLINE            99310773
PUBMED            10381312
                2 (bases 1 to 783)
REFERENCE          Rebholz,W.E.R. and Harley,E.H.
AUTHORS           Direct Submission
TITLE              Submitted (21-OCT-1997) Chemical Pathology, University of Cape Town
JOURNAL            Medical School, Observatory, Cape Town 7925, South Africa
FEATURES
source            1..783
                  /organism="Raphicerus campestris"
                  /organelle="mitochondrion"
                  /db_xref="taxon:59544"
                  /map="8970-9752 relative to the cow (Bos taurus) sequence"
                  /note="From cell culture"
                  1..>783
                  /gene="cox iii"
                  1..>783
CDS

```

BASE COUNT	ORIGIN	205 a	223 c	114 g	241 t
Query Match	30.9%	Score 35.2;	DB 4;	Length 783;	
Best Local Similarity	58.7%	Pred. No. 0.76;			
Matches	61; Conservative	0;	Mismatches 43;	Indels 0;	Gaps 0;
Qy	8	acggcagaccgaagaanaacccgagattgacctagagattggccatcgccctgctgacct	67		
Db	458	AAGGCAATCGCTACCCACTACTTCAACGCTTTATTTATACCAATTCAGCGATGACGCTGATCT	517		
Qy	68	ttaccctgtgccaaagcctatgaatatatgaataatcgctccatcc	111		
Db	518	TTACACTGCTCCAGACGCTCAGATATTTATGAGACACCCCTTTACC	561		
RESULT 14	AF030472	783 bp	DNA	MAM	26-JUL-1999
LOCUS	Gazella rufifrons cytochrome oxidase subunit III (cox iii) gene,				
DEFINITION	mitochondrial gene encoding mitochondrial protein, partial cds.				
ACCESSION	AF030472				
VERSION	AF030472.1	GI:2731921			
KEYWORDS	red-fronted gazelle				
SOURCE	Mitochondrial Gazella rufifrons				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;				
	Bovidae; Antilopinae; Gazella.				
REFERENCE	1 (bases 1 to 783)				
AUTHORS	Rehbolz,W. and Harley,E.				
TITLE	Phylogenetic relationships in the bovid subfamily Antilopinae based				
JOURNAL	on mitochondrial DNA sequences				
MEDLINE	Mol. Phylogenet. Evol. 12 (2), 87-94 (1999)				
PUBMED	99310773				
REFERENCE	2 (bases 1 to 783)				
AUTHORS	Rehbolz,W.E.R. and Harley,E.H.				
TITLE	Direct Submission				
JOURNAL	Submitted (21-OCT-1997) Chemical Pathology, University of Cape Town				
FEATURES	Medical School, Observatory, Cape Town 7925, South Africa				
source	Location/Qualifiers				
	1. .783				
	/organism="Gazella rufifrons"				
	/organelle="mitochondrion"				
	/db_xref="taxon:69304"				
	/map="8970-9752 relative to the cow (Bos taurus) sequence"				
	/note="From blood sample"				
	1. .>783				
	/gene="cox iii"				
	1. .>783				
	/gene="cox iii"				
	/codon_start=1				
	/transl_table=2				
	/product="cytochrome oxidase subunit III"				
	/protein_id="AAB93601.1"				
	/db_xref="GI:2731921"				
	/translation="MTROTTHAYHMVNPSPFWPLTGALISALLMTSGILMWFPHNSTILIM				
	IGLTNNLTMYQWMDRVIRESTFOGHHTPVQKGLRGMILFIISEVLPFGTFNAFY				
	HSSIAAPPELGCCWPTGIVRISLPLEVPLNTSVLASGSITVAHSHLMGNNKRYPM				
	QALFTIATGVFTLLQASVYEAPEFTISDGVSTFPVATGFGHIVIIIGSTFLIVC				
	FEROLKHEFTSNHHGFPEANAMVHFVYVWLFYISLYIMWGS"				
gene					
CDS					

Query Match	30.5%;	Score 34.8;	DB 4;	Length 783;
Best Local Similarity	57.3%;	Pred. No. 1;		
Matches 63; Conservative	0;	Mismatches 47;	Indels 0;	Gaps 0;

RESULT	15
AF030473	
LOCUS	763 bp DNA
DEFINITION	Gazella thomsoni cytochrome oxidase subunit III (cox iii) gene
ACCESSION	AF030473
VERSION	AF030473.1
KEYWORDS	GI:2731923

REFERENCE	AUTHORS	TITLE
1 (bases 1 to 783)	Rebholz, W. and Harley, E.	Phylogenetic relationships in the bovid subfamily Bovinae

on mitochondrial DNA sequences
Mol. Phylogenet. Evol. 12 (2), 87-94 (1999)
MEDLINE 99310773

PUBMED
 10381312
 2 (bases 1 to 783)
 Rehholz, W.E. R. and Harley, E. H.
 Direct Submission
 Submitted (21-OCT-1997) Chemical Pathology, University of Cape Town
 Medical School, Observatory, Cape Town 7925, South Africa
 Location/Qualifiers

Source

```
gene      1. .>783
          /gene="cox 11"
CDS       1. .>783
```

BASE COUNT	ORIGIN
217 a	238 c
107 g	221 t

Query Match	30.58;	Score	34.8;	DB	4;	Length	783;
Best Local Alignment	30.58;	Score	34.8;	DB	4;	Length	783;

Best Local Similarity	57.38;	DB 4;	Length 783;
Matches	63; Conservative	0;	Mismatches 47; Indels 0; Gaps 0

2 tcgtccacgycgacgcagaaacgcgattgctctagcgattgcacatcgcccttgct 67

Db 452 TCATAGAGGAATATCCCAACCATCTTCTCAAGGCGCTATTTCATTACCATCGACATGAGGG 511

Qy 62 ggaatctttaccctgtgcgaagcctatgaatatatgaatgcgtccatcc 111

Db 512 TATACTTCACACTACTCTCCAGCGCTCAGATATATTAGCAGAGACCGCTTTTAC 561

Search completed: January 10, 2002, 00:56:35
Job time: 3083 sec

Patent No. 5840695
GENERAL INFORMATION:
APPLICANT: FRANK, GLENN R.
APPLICANT: HUNTER, SHIRLEY WU
APPLICANT: WALLENFELS, LYNDY
TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross P.C.
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: U.S.A.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US-08/630,822A
FILING DATE: 11-APR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: CONNELL, GARY J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-17-C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 60:
SEQUENCE CHARACTERISTICS:
LENGTH: 157 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-630-822A-60

alignment_scores:
Quality: 53.00 Length: 27
Ratio: 2.944 Gaps: 1
Percent Similarity: 66.667 Percent Identity: 40.741

alignment_block:
US-09-712-768-5/rev x US-08-630-822A-60 ..
Align seg 1/1 to: US-08-630-822A-60 from: 1 to: 157

80 TGGCACAGGGTAAGATCCAGCCAGCGCATGCGATCGTAGGCCAAT 31
||| : : : : ||||| : : : : |||||
33 TTPASPLYSLEUPROPHETVALPROARGPROSERLEUVALAIAASPROLL 49
30 CGCGGTTTCTTGGCGTCGCCGTGGAGCA 2
| : : : : ||| ||| : : : : |||
49 ethralaphe.CyslysarglysPROARG 58

seq_name: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:US-09-005-069-60

seq_documentation_block:
Sequence 60, Application US/09005069
Patent No. 5932470
GENERAL INFORMATION:
APPLICANT: FRANK, GLENN R.
APPLICANT: HUNTER, SHIRLEY WU
APPLICANT: WALLENFELS, LYNDY
TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross P.C.
STREET: 1700 Lincoln Street, Suite 3500

CITY: Denver
STATE: Colorado
COUNTRY: U.S.A.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/005,069
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/630,822
FILING DATE: 11-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: CONNELL, GARY J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-17-C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 60:
SEQUENCE CHARACTERISTICS:
LENGTH: 157 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-005-069-60

alignment_scores:
Quality: 53.00 Length: 27
Ratio: 2.944 Gaps: 1
Percent Similarity: 66.667 Percent Identity: 40.741

alignment_block:
US-09-712-768-5/rev x US-09-005-069-60 ..
Align seg 1/1 to: US-09-005-069-60 from: 1 to: 157

80 TGGCACAGGGTAAGATCCAGCCAGCGCATGCGATCGTAGGCCAAT 31
||| : : : : ||||| : : : : |||||
33 TTPASPLYSLEUPROPHETVALPROARGPROSERLEUVALAIAASPROLL 49
30 CGCGGTTTCTTGGCGTCGCCGTGGAGCA 2
| : : : : ||| ||| : : : : |||
49 ethralaphe.CyslysarglysPROARG 58

seq_name: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:US-09-029-213B-22

seq_documentation_block:
Sequence 22, Application US/09029213B
Patent No. 6180098
GENERAL INFORMATION:
APPLICANT: CHRISTIAN, Peter D.
TITLE OF INVENTION: RECOMBINANT HELICOVERA BACULOVIRUSES
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: McDermott, Will & Emery
STREET: 600 13th Street, NW
CITY: Washington
STATE: District of Columbia
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

Page 3

```

? FILING DATE: 13-NOV-1992
? ATTORNEY/AGENT INFORMATION:
? NAME: Hamburg, C Bruce
? REGISTRATION NUMBER: 22,389
? REFERENCE/DOCKET NUMBER: F-4551
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (212) 986-7340
? TELEFAX: (212) 953-7733
? INFORMATION FOR SEQ ID NO: 3:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1239 amino acids
? TYPE: amino acid
? STRANDEDNESS: single strand
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? ORIGINAL SOURCE:
? ORGANISM: mouse
? TISSUE TYPE: brain
? PUBLICATION INFORMATION:
? AUTHORS: Mesayoshi MISHINA
? TITLE: NOVEL PROTEIN AND GENES CODING THE SAME
? RELEVANT RESIDUES IN SEQ ID NO: 3: FROM 1 TO 1239
?
? US-08-026-138E-3
?
? alignment_scores:
? Quality: 50.50 Length: 43
? Ratio: 2.806 Gaps: 2
? Percent Similarity: 41.660 Percent Identity: 32.558
?
? alignment_block:
? US-09-712-768-5 x US-08-026-138E-3 ..
?
? Align seg 1/1 to: US-08-026-138E-3 from: 1 to: 1239
?
? 3 CGTCCAGCGCGACGCCAAGAAAACCGCATGGCCT..... 38
? ||| ||||| |||::||| |||
? 1007 ArgThrSerArgProGluArgAlaValArgAlaLeuProGluArgSerLeu 1023
? 39 ....AGCGATTGCCAT.....C 51
? ::| |||||
? 1023 UHSAIAIHISYSHITFTYSerSerPheProArgAlaGluArgSerGly 1040
? 52 GGCCTGCGTGGATCTTTACCTGTGCCA 80
? ||||::||| ||::||| |||
? 1040 rgrProPheLeuProLeuPheProGluPro 1049
?
? seq_name: /cgn2_6/prodata/2/1aa/6B.COMB.pep:US-09-159-106-13
?
? seq_documentation_block:
? Sequence 13, Application US/09159106
? Patent No. 6284509
? GENERAL INFORMATION:
? APPLICANT: Ferrer, Pau
? APPLICANT: Diers, Ivan
? APPLICANT: Hedegaard, Lisbeth
? APPLICANT: Halkier, Torben
? TITLE OF INVENTION: An Enzyme With -1,3-Glucanase
? TITLE OF INVENTION: Activity
? FILE REFERENCE: 4693.204-US
? CURRENT APPLICATION NUMBER: US/09/159,106
? CURRENT FILING DATE: 1998-09-23
? EARLIER APPLICATION NUMBER: 0427/96
? EARLIER FILING DATE: 1996-12-04
? EARLIER APPLICATION NUMBER: 0885/96
? EARLIER FILING DATE: 1996-08-23
? EARLIER APPLICATION NUMBER: PCT/DK97/00160
? EARLIER FILING DATE: 1997-04-14
? NUMBER OF SEQ ID NOS: 15
? SOFTWARE: FastSeq for Windows Version 3.0
? SEQ ID NO 13
? LENGTH: 303
? TYPE: PRT

```

Align seg 1/1 to: US-09-159-106-13

```

; Sequence 11, Application US/09159106
; Patent NO. 6284509
; GENERAL INFORMATION

```

alignment_scores:

US-09-159-106-11 x US-09-159-106-11

8 ACGCGACACCCAGAAAACCGCGTGTGGCGTTCGCAATGCCATCGGCGCTT 57
14 ThrProThrGlyAlaGlyGlyProAlaAlaArgArgLeuAlaSerAlaAla 30
58 GCGTCGACATCTTTACCGCTGTGCC 79
30 uValAlaAlaLeuThrAlaAla 37

4

alignment_scores:

PCI-0591-09422-19

[illegible]

```
seq_name: /cgn2_6/prodata/2/1aa/5B.COMB.pep:US-08-484-101B-36
seq_documentation_block:
; Sequence 36, Application US/08484101B
; Patent No. 5824868
; GENERAL INFORMATION:
```


Thu Jan 10 08:15:40 2002

us-09-712-768-5.ra1

Page 5

```

APPLICANT: California Institute of Technology
TITLE OF INVENTION: PLANTS HAVING MODIFIED RESPONSE TO ETHYLENE
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESS: Richard F. Treccartin
STREET: 3400 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,101B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/
FILING DATE: 01-JUL-1994
CLASSIFICATION: 800
PRIOR APPLICATION NUMBER: US 08/086,555
FILING DATE: 01-JUL-1993
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Treccartin, Richard F.
REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: A-57515-2/RFT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 635 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-101B-36

alignment_scores:
Quality: 49.00 Length: 28
Ratio: 2.579 Gaps: 2
Percent Similarity: 67.857 Percent Identity: 39.286

alignment_block:
US-09-712-768-5 x US-08-484-101B-36 ..
Align seg 1/1 to: US-08-484-101B-36 from: 1 to: 635
2 TCGTCACGCGGACCGCAGAAACCGGATGGCTAGCGATTGCCATC 51
||| |||:||||| |||:||||| |||:|||||
480 SerIleGluAlaSerValAlaIysPro.GluTyrAlaArgAspCysHisP 496
52 GGCCTTGCGTGGATCTTACCTGTGCCAAGC 83
||| |||:||||| |||:||||| |||:|||||
496 roPro.....GluMetPheProMetProSer 504

seq_name: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:US-08-484-101B-50
seq_documentation_block:
Sequence 50, Application US/08484101B
Patent No. 5824868
GENERAL INFORMATION:
APPLICANT: California Institute of Technology
TITLE OF INVENTION: PLANTS HAVING MODIFIED RESPONSE TO ETHYLENE
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:

```

```

ADDRESS: Richard F. Treccartin
STREET: 3400 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,101B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/
FILING DATE: 01-JUL-1994
CLASSIFICATION: 800
PRIOR APPLICATION NUMBER: US 08/086,555
FILING DATE: 01-JUL-1993
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Treccartin, Richard F.
REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: A-57515-2/RFT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 635 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-101B-50

alignment_scores:
Quality: 49.00 Length: 28
Ratio: 2.579 Gaps: 2
Percent Similarity: 67.857 Percent Identity: 39.286

alignment_block:
US-09-712-768-5 x US-08-484-101B-50 ..
Align seg 1/1 to: US-08-484-101B-50 from: 1 to: 635
2 TCGTCACGCGGACCGCAGAAACCGGATGGCTAGCGATTGCCATC 51
||| |||:||||| |||:||||| |||:|||||
480 SerIleGluAlaSerValAlaIysPro.GluTyrAlaArgAspCysHisP 496
52 GGCCTTGCGTGGATCTTACCTGTGCCAAGC 83
||| |||:||||| |||:||||| |||:|||||
496 roPro.....GluMetPheProMetProSer 504

seq_name: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:US-08-714-524D-36
seq_documentation_block:
Sequence 36, Application US/08714524D
Patent No. 6294716
GENERAL INFORMATION:
APPLICANT: Chang, Caren
APPLICANT: Blecker, Anthony B
TITLE OF INVENTION: PLANTS HAVING MODIFIED RESPONSE TO ETHYLENE
FILE REFERENCE: A-57515-4
CURRENT APPLICATION NUMBER: US/08/714,524D
CURRENT FILING DATE: 1996-09-16
NUMBER OF SEQ ID NOS: 56
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 36

```

Thu Jan 10 08:15:40 2002

us-09-712-768-5.ra1

```

; LENGTH: 635
; TYPE: PRT
; ORGANISM: Lycopersicon esculentum
US-08-714-524D-36

alignment_scores:
  Quality: 49.00      Length: 28
  Ratio: 2.579        Gaps: 2
  Percent Similarity: 67.857  Percent Identity: 39.286
alignment_block:
US-09-712-768-5 x US-08-714-524D-36 ..
Align seg 1/1 to: US-08-714-524D-36 from: 1 to: 635
2 TCCTCCACGGCGACCGCAAGAAACCGGATGGCTAGCGATGGCCATC 51
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
480 SerIleGluAlaSerValAlaLysPro.GluTyrAlaIaArgAspCysHisP 496
52 GGCCTTGCTGGATCTTACCGCTGTGCCAAGC 83
|||:|||||:|||||:|||||:|||||:|||||:|||||:
496 roPro.....GluMetPheProMetProSer 504

seq_name: /cgn2_6/ptodata/2/1aa/6B-COMB.pep:US-08-714-524D-50
seq_documentation_block:
; Sequence 50, Application US/08714524D
; Patent No. 6294716
; GENERAL INFORMATION:
; APPLICANT: Meyerowitz, Elliott M
; APPLICANT: Chang, Caren
; TITLE OF INVENTION: PLANTS HAVING MODIFIED RESPONSE TO ETHYLENE
; FILE REFERENCE: a-57515-4
; CURRENT APPLICATION NUMBER: US/08/714,524D
; CURRENT FILING DATE: 1996-09-16
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 50
; LENGTH: 635
; TYPE: PRT
; ORGANISM: Lycopersicon esculentum
US-08-714-524D-50

alignment_scores:
  Quality: 49.00      Length: 28
  Ratio: 2.579        Gaps: 2
  Percent Similarity: 67.857  Percent Identity: 39.286
alignment_block:
US-09-712-768-5 x US-08-714-524D-50 ..
Align seg 1/1 to: US-08-714-524D-50 from: 1 to: 635
2 TCCTCCACGGCGACCGCAAGAAACCGGATGGCTAGCGATGGCCATC 51
|||:|||||:|||||:|||||:|||||:|||||:|||||:
480 SerIleGluAlaSerValAlaLysPro.GluTyrAlaIaArgAspCysHisP 496
52 GGCCTTGCTGGATCTTACCGCTGTGCCAAGC 83
|||:|||||:|||||:|||||:|||||:|||||:|||||:
496 roPro.....GluMetPheProMetProSer 504

seq_name: /cgn2_6/ptodata/2/1aa/6B-COMB.pep:US-08-460-269C-2
seq_documentation_block:
; Sequence 2, Application US/08460269C
; Patent No. 6197548
; GENERAL INFORMATION:
; APPLICANT: CLARE, JEFFREY J.
; APPLICANT: ROMANOS, MICHAEL A.
; TITLE OF INVENTION: EXPRESSION OF HETEROLOGOUS PROTEIN IN
```

```

; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Millen, White, Zelano & Branigan, P.C.
; STREET: 2200 Clarendon Blvd., Suite 1400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,269C
; FILING DATE: 02-Jun-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lebowitz, Richard M.
; REGISTRATION NUMBER: 37,067
; REFERENCE/DOCKET NUMBER: Popov-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 243-6333
; TELEFAX: (703) 243-6410
; INFORMATION FOR SEQ ID NO: 2:
; LENGTH: 910 amino acids
; TYPE: amino acid
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-08-460-269C-2

alignment_scores:
  Quality: 49.00      Length: 30
  Ratio: 2.450        Gaps: 1
  Percent Similarity: 66.667  Percent Identity: 36.667
alignment_block:
US-09-712-768-5 x US-08-460-269C-2 ..
Align seg 1/1 to: US-08-460-269C-2 from: 1 to: 910
4 GTCCAGCGGCGACCGCAAGAAACCGGATGGCTAGCGATGGCCATCG 53
|||:|||||:|||||:|||||:|||||:|||||:|||||:
863 LeuArgGlyThrArgAlaGluLeuGlyLeuGlyMetAlaAlaLeuG 879
54 CCTGGCTGGATCTTACCGCTGTGCCAAGCTTGAATAT 93
|||:|||||:|||||:|||||:|||||:|||||:
879 yArgGly.....HisSerLeuTyrAlaSerTyrGluTyr 890

seq_name: /cgn2_6/ptodata/2/1aa/6B-COMB.pep:US-08-460-269C-4
seq_documentation_block:
; Sequence 4, Application US/08460269C
; Patent No. 6197548
; GENERAL INFORMATION:
; APPLICANT: CLARE, JEFFREY J.
; APPLICANT: ROMANOS, MICHAEL A.
; TITLE OF INVENTION: EXPRESSION OF HETEROLOGOUS PROTEIN IN
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Millen, White, Zelano & Branigan, P.C.
; STREET: 2200 Clarendon Blvd., Suite 1400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
;
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,269C
; FILING DATE: 02-Jun-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lebovitz, Richard M.
; REGISTRATION NUMBER: 37,067
; REFERENCE/DOCKET NUMBER: Popov-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 243-6333
; TELEFAX: (703) 243-6410
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 911 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-08-460-269C-4

alignment_scores:
      Quality: 49.00      Length: 30
      Ratio: 2.450      Gaps: 1
Percent Similarity: 66.667      Percent Identity: 36.667

alignment_block:
US-09-712-768-5 x US-08-460-269C-4  ..

Align seg 1/1 to: US-08-460-269C-4 from: 1 to: 911

      4 GTCCACGGCGACCGCAAGAAACCGCATTTGGCTAGCGATGCCATCGG 53
      :::::| | | | | :::::| | | | | | | | | | | | | | |
      864 LeuArgGlyThrArgAlaGluLeuGlyLeuGlyMetAlaAlaLeuGI 880
      54 CTTGGCTGGATCTTTACCGTGTGCCAAGCCTATGAATAT 93
      | | | | | :::::| | | | | | | | | | | | | | |
      880 YArgGly.....HisSerLeuTyrAlaSerTyrGluTyr 891

seq_name: /cgn2.6/prodata/2/laa/5B_COMB.pep:US-08-460-269C-6

seq_documentation_block:
; Sequence 6, Application US/08460269C
; Patent No. 6197548
; GENERAL INFORMATION:
; APPLICANT: CLARE, JEFFREY J.
; ROMANOS, MICHAEL A.
; TITLE OF INVENTION: EXPRESSION OF HETEROLOGOUS PROTEIN IN
; YEAST
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Millen, White, Zelano & Branigan, P.C.
; STREET: 2200 Clarendon Blvd., Suite 1400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22201
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,269C
; FILING DATE: 02-Jun-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lebovitz, Richard M.
; REGISTRATION NUMBER: 37,067
; REFERENCE/DOCKET NUMBER: Popov-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 243-6333
; TELEFAX: (703) 243-6410
```

```
;
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 922 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-08-460-269C-6

alignment_scores:
      Quality: 49.00      Length: 30
      Ratio: 2.450      Gaps: 1
Percent Similarity: 66.667      Percent Identity: 36.667

alignment_block:
US-09-712-768-5 x US-08-460-269C-6  ..

Align seg 1/1 to: US-08-460-269C-6 from: 1 to: 922

      4 GTCCACGGCGACCGCAAGAAACCGCATTTGGCTAGCGATGCCATCGG 53
      :::::| | | | | :::::| | | | | | | | | | | | | | |
      875 LeuArgGlyThrArgAlaGluLeuGlyLeuGlyMetAlaAlaLeuGI 891
      54 CTTGGCTGGATCTTTACCGTGTGCCAAGCCTATGAATAT 93
      | | | | | :::::| | | | | | | | | | | | | | |
      891 YArgGly.....HisSerLeuTyrAlaSerTyrGluTyr 902
```



```

|||||
1 1LevalHisGlyAspArgLysThrAlaIleGlyLeuAlaIleAlaIle 17
51 CGGCTTGCGTGCATCTTACCCGTGCTCCAGCCTATGAAATATTATGAAA 100
17 eGlyLeuGlyLysThrPheThrLeuGlyGlnAlaIleTyrGlyTyrGlyLys 34
101 TCGTCATACCGCAA 114
|||||
34 LevalHisThrGlu 38

```

seq_name: /SIDS2/gcgdata/geneseq/geneseqp/AA2001.DAT: AAB63840

seq_documentation_block:

ID AAB63840 standard; Protein: 93 AA.

AC AAB63840;

26-MAR-2001 (first entry)

DE Human prostate cancer associated antigen protein sequence SEQ ID NO: 1202.

XX Human; breast cancer; gastric cancer; prostate cancer; diagnosis;

XX cancer associated antigen; cytostatic; cancer vaccine.

XX Homo sapiens.

OS WO200073801-A2.

PN 07-DEC-2000.

PD 26-MAY-2000; 2000WO-US14749.

PF 28-MAY-1999; 99US-0136526.

PR 10-SEP-1999; 99US-0153454.

XX (LUDW-) LUDWIG INST CANCER RES.

PA

PI Obata Y;

XX WPI: 2001-025274/03.

DR

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

```

1 ATGTCACAGCGACCGACAGAAAACCGCATGTGCGCTAGCGATGCCAT 50
5 LeuIleGluAsnAsnArgAsnGlnIleIleGlnAlaLeuLeuIleThr 21
51 CGGCTTGCGTGCATCTTACCCGTGCTCCAGCCTATGAAATATTATGAAA 99
21 eLeuLeuGlyLeuTyrPheThrLeuGlnAlaSerGlyTyr-PheGlu 37

```

seq_name: /SIDS2/gcgdata/geneseq/geneseqp/AA2000.DAT: AAB29910

seq_documentation_block:

ID AAB29910 standard; Protein: 133 AA.

AC AAB29910;

09-FEB-2001 (first entry)

DE Human secreted protein BLAST search protein SEQ ID NO: 168.

XX Cytostatic; immunosuppressive; nocotropic; neuroprotective; antiviral;

XX anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; anticancer;

XX vanguard; anticonvulsant; antibacterial; antifungal; antiparasitic;

XX cardiatic; gene therapy; cancer; immune disorder; cardiovascular disorder;

XX neurological disease; infection; human; secreted protein.

XX Homo sapiens.

OS WO200061779-A1.

PN 19-OCT-2000.

PD 06-APR-2000; 2000WO-US09068.

PF 09-APR-1999; 99US-0128699.

PR 20-JAN-2000; 2000US-0177050.

XX (HUMA-) HUMAN GENOME SCI INC.

PA

PI Rosen CA, Ruben SM, Komatsoulis G;

XX WPI: 2000-647424/62.

DR

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Sequence 133 AA;

alignment_scores:

Quality: 62.00

Ratio: 2.583

Length: 33

Gaps: 0

alignment_scores:
Quality: 62.00
Ratio: 2.583
Percent Similarity: 72.727
Percent Identity: 42.424

alignment_block:
US-09-712-768-5 x AAB63840

Align seg 1/1 to: AAB63840 from: 1 to: 93

Percent Similarity: 72.727 Percent Identity: 42.424

alignment_block:

US-09-712-768-5 x AAB75604 ..

Align seg 1/1 to: AAB75604 from: 1 to: 133

```

1 ATGCTCCAGCGCGCAAGAAACCGCGATTGGCTTAGCGATTGCCAT 50
   : : : : : : : : : : : : : : : : : : : : : : : :
23 LeuLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 39
   : : : : : : : : : : : : : : : : : : : : : : :
51 CGGCGTTGGCTGATCTTACCTGTCGCAAGCCTATGATATATGAA 99
   | | | | | | | | | | | | | | | | | | | | | | |
39 eLeuLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 55

```

seq.name: /SID52/gcgdata/geneseq/geneseq/AA2001.DAT: AAB75604

seq_documentation_block:

ID AAB75604 standard; Protein: 203 AA.

AC AAB75604;

DT 06-APR-2001 (first entry)

Human secreted protein sequence encoded by gene 46 SEQ ID NO:158.

Human: secreted protein; immunosuppressive; antirheumatic; antiproliferative; cytostatic; cardiac; vasotropic; cerebroprotective; neurotrophic; neuroprotective; antibacterial; virucide; fungicide; opthalmological; vulnary; autoimmune disease; cardiovascular disorder; hyperproliferative disorder; cerebrovascular disorder; wound healing; nervous system disorder; ocular disorder; skin ageing; chemotaxis; food additive.

OS Homo sapiens.

PN WO200077026-A1.

PD 21-DEC-2000.

PF 01-JUN-2000; 2000WO-US14973.

PR 11-JUN-1999; 99US-0138630.

PA (HUMA-) HUMAN GENOME SCI INC.

PA (ROSE/) ROSEN C A.

PI Rosen CA, Ruben SM, Komatsoulis GA;

PI WPI: 2001-071258/08.

DR N-PSDB; AAF64221.

PT Nucleic acid molecules encoding human secreted proteins, used in preventing, treating or ameliorating a disorder, e.g. Alzheimer's and Parkinson's diseases and cancers -

PS Disclosure: page 94; 542pp; English.

Human secreted proteins AAB75506 - AAB75554 are encoded by polynucleotide sequences AAF64176 - AAF64224. The specification includes amino acid sequences AAB75555 - AAB75606 which represent fragments of the human secreted proteins, and protein sequences with which they share homology. The proteins and polynucleotides, their agonists and antagonists have activities dependent on the tissues and cells in which they are expressed, examples of these activities include: immunosuppressive; antirheumatic; antiproliferative; cytostatic; cardiac; vasotropic; cerebroprotective; neurotrophic; neuroprotective; antibacterial; antiproliferative; opthalmological; and vulnary. The proteins, polynucleotides, agonists and antagonists can be used to treat or detect diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,

anglogenesis, nervous system disorders e.g. Alzheimer's disease, infections caused by bacteria, viruses and fungi can be used to aid wound healing and epithelial cell proliferation, to prevent skin ageing due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities. Included in the invention are polynucleotide sequences AAF64167 - AAF64175 and peptide AAB75505 which are used in the isolation, identification and characterisation of the proteins of the invention.

XX Sequence 203 AA;

alignment_scores: Length: 33
Quality: 62.00
Ratio: 2.583
Percent Similarity: 72.727 Percent Identity: 42.424

alignment_block:

US-09-712-768-5 x AAB75604 ..

Align seg 1/1 to: AAB75604 from: 1 to: 203

```

1 ATGCTCCAGCGCGCAAGAAACCGCGATTGGCTTAGCGATTGCCAT 50
   : : : : : : : : : : : : : : : : : : : : : : :
93 LeuLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 109
   : : : : : : : : : : : : : : : : : : : : : : :
51 CGGCGTTGGCTGATCTTACCTGTCGCAAGCCTATGATATATGAA 99
   | | | | | | | | | | | | | | | | | | | | | | |
109 eLeuLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 125

```

seq.name: /SID52/gcgdata/geneseq/geneseq/AA1999.DAT: AAY35610

seq_documentation_block:

ID AAY35610 standard; Protein: 455 AA.

AC AAY35610;

DT 13-SEP-1999 (first entry)

Human secreted protein sequence.

Human: secreted protein; pneumonia; bronchitis; heart disease; sarcoidosis; respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis; sinusitis; purulent otitis media; erythema nodosum; pharyngitis; vaccine; neutralising epitope.

OS Chlamydia pneumoniae.

PN WO9927105-A2.

PD 03-JUN-1999.

PF 20-NOV-1998; 98WO-IB01890.

PR 04-NOV-1998; 98US-0107078.

PR 21-NOV-1997; 97FR-0014673.

PA (GEST) GENSET.

PA Griffais R;

PI WPI: 1999-357842/30.

PT Genome sequence of Chlamydia pneumoniae

PS Page 1338-1339; Disclosure: 1912pp; English.

AA354584-Y35879 represent the proteins encoded by all the open reading frames in the complete genome (see AA354584) of Chlamydia pneumoniae. C. pneumoniae causes respiratory disease such as pneumonia and bronchitis and is thought to be a contributing factor in heart

CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema
CC nodosum or pharyngitis. The polypeptides encoded by the open reading
CC frames of the C. pneumoniae genome (see AAY34584-Y35879) can be used in
CC immunogenic compositions as vaccines. Vectors containing C. pneumoniae
CC nucleotide sequences can also be used as immunogenic compositions,
CC especially where the vector directs the expression of a neutralising
CC epitope of C. pneumoniae.
XX
SQ Sequence 455 AA.

alignment_scores:

Quality: 62.00 Length: 18
Ratio: 4.133 Gaps: 0
Percent Similarity: 83.333 Percent Identity: 50.000

alignment_block:

US-09-712-768-5 x AAY35610 ..

Align seg 1/1 to: AAY35610 from: 1 to: 455

61 TGGATCTTACCCCTGCGCAAGCCATGATATATGAAATCGTCATAC 110
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
35 TrpleuTyTrhleucYcglInserTrleuHisLysLysLeuValPhepr 51

111 CGAA 114
|||

51 oGlu 52

seq_name: /SID2/gcgdata/geneseq/geneseq/AA2000.DAT: AAB29909
seq_documentation_block:

ID AAB29909 standard; Protein: 133 AA.

XX AAB29909;

DT 09-FEB-2001 (first entry)

DE Human secreted protein BLAST search protein SEQ ID NO: 167.

XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
XX anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; antitumor;
XX vulnerability; anticonvulsant; antibacterial; antifungal; antiparasitic;
XX cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;
XX neurological disease; infection; human; secreted protein.
XX Homo sapiens.

PN WO200061779-A1.

PD 19-OCT-2000.

PE 06-APR-2000; 2000WO-US09068.

PR 09-APR-1999; 99US-0128699.

PR 20-JAN-2000; 2000US-0177030.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM, Komatsoulis G;

DR WPI: 2000-647424/62.

XX Isolated nucleic acid molecule encoding a human secreted protein is
XX used in preventing, treating or ameliorating a medical condition -
XX
XX Disclosure: Page 487; 495pp; English.

CC The invention relates to the isolation of genes AAC63410-C63458 encoding
CC the human secreted proteins AAB29802-B29850. This sequence represents a
CC peptide fragment homologous to the protein encoded by the gene given
CC in the descriptor line. The sequence is a search result from a BLASTX
CC homology search. The genes and proteins are useful for preventing,

CC ameliorating or treating medical conditions, e.g. by protein or gene
CC therapy. The genes are isolated from a range of human tissues disclosed
CC in the specification. The nucleic acids, proteins, antibodies and
CC (ant)agonists are useful in the diagnosis, treatment and prevention of:
CC (a) cancer, e.g. breast and ovarian cancer, and other cancers of the
CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,
CC lung, or urogenital; (b) immune disorders e.g. Addison's disease,
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; (c) cardiovascular disorders such
CC as myocardial ischaemia; (d) wound healing; (e) neurological diseases
CC e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such
CC as viral, bacterial, fungal and parasitic infections.
XX
SQ Sequence 133 AA.

alignment_scores:

Quality: 60.00 Length: 33
Ratio: 2.500 Gaps: 0
Percent Similarity: 72.727 Percent Identity: 42.424

alignment_block:

US-09-712-768-5 x AAB29909 ..

Align seg 1/1 to: AAB29909 from: 1 to: 133

1 ATGCTCCACGGCGCGCAAGAAACCGCGATTGGCTACGATGCCAT 50
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

23 LeuMetGluAsnAsnArgAsnGlnMetCLeuAlaLeuLeuLeuThrIle 39

51 CGGCTTGCGCTGGATCTTACCTGTGCGCAAGCTATGAAATATATGAA 99

1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

39 eLeuLeuGlyLeuTyTrPheTrhleucGlnAlaSerGluTyTrPheGlu 55

seq_name: /SID2/gcgdata/geneseq/geneseq/AA2000.DAT: AAB36196

seq_documentation_block:

ID AAB36196 standard; Protein: 172 AA.

XX AAB36196;

DT 02-MAR-2001 (first entry)

DE Cowdria ruminantium 4hwoF1 polypeptide.

XX Cowdria ruminantium; MAP1; major antigenic protein 1; antirickettsial;
XX KW vaccine; gene therapy; Rickettsia; Cowdria; Anaplasma; map2; ihwoF3;
XX 4hwoF1; 18hwoF1; 3gdorF3.

OS Cowdria ruminantium.

PN WO200065063-A2.

PD 02-NOV-2000.

PE 21-APR-2000; 2000WO-US10886.

PR 22-APR-1999; 99US-0130725.

PA (UYFL) UNIV FLORIDA.

PI Barbet AF, Bowie MV, Ganta RR, Burridge MJ, Mahan SM, McGuire TC;
PI Rurangirwa FR, Moreland AL, Simbi BH, Whitmore MW, Alleman AR;

DR WPI: 2000-679675/66.

DR N-PSDB; AAC68713.

XX New polynucleotides useful as DNA vaccines for conferring immunity to
XX rickettsial infection e.g. heartwater caused by Cowdria ruminantium,
XX comprises major antigenic protein 1 or 2 gene of rickettsial pathogens

XX PD 07-FEB-2001.

AA
DR WPI; 2001-071258/08.


```

PR 29-APR-1998; 98US-0083545.
PR 29-APR-1998; 98US-0083554.
PR 29-APR-1998; 98US-0083558.
PR 29-APR-1998; 98US-0083559.
PR 30-APR-1998; 98US-0083742.
PR 05-MAY-1998; 98US-0084366.
PR 06-MAY-1998; 98US-0084414.
PR 06-MAY-1998; 98US-0084441.
PR 07-MAY-1998; 98US-0084598.
PR 07-MAY-1998; 98US-0084600.
PR 07-MAY-1998; 98US-0084627.
PR 07-MAY-1998; 98US-0084637.
PR 07-MAY-1998; 98US-0084639.
PR 07-MAY-1998; 98US-0084640.
PR 07-MAY-1998; 98US-0084643.
PR 13-MAY-1998; 98US-0085323.
PR 13-MAY-1998; 98US-0085338.
PR 13-MAY-1998; 98US-0085339.
PR 15-MAY-1998; 98US-0085573.
PR 15-MAY-1998; 98US-0085579.
PR 15-MAY-1998; 98US-0085580.
PR 15-MAY-1998; 98US-0085582.
PR 15-MAY-1998; 98US-0085689.
PR 15-MAY-1998; 98US-0085697.
PR 15-MAY-1998; 98US-0085700.
PR 15-MAY-1998; 98US-0085704.
PR 18-MAY-1998; 98US-0086023.
PR 22-MAY-1998; 98US-0086392.
PR 22-MAY-1998; 98US-0086414.
PR 22-MAY-1998; 98US-0086430.
PR 22-MAY-1998; 98US-0086486.
PR 28-MAY-1998; 98US-0087098.
PR 28-MAY-1998; 98US-0087106.
PR 28-MAY-1998; 98US-0087208.
PR 30-JUL-1998; 98US-0094651.
PR 11-SEP-1998; 98US-0100038.
XX
XX (GETH ) GENENTECH INC.
XX
XX Wood WI, Goddard A, Gurney A, Yuan J, Baker KP, Chen J;
XX
XX WPI: 1999-551358/46.
XX N-PSDB; AA234275.
XX
XX New secreted and transmembrane polypeptides and their polynucleotides,
XX useful for treating blood coagulation disorders, cancers and cellular
XX adhesion disorders -
XX
XX Claim 12; Fig 190; 530pp; English.
XX
XX The present invention describes secreted and transmembrane polypeptides
XX and their polynucleotides. The nucleotide sequences are useful as
XX sources of probes, primers, for chromosome mapping, and for generation
XX of antisense sequences. They can also be used to create transgenic
XX animals. The proteins can be used to treat a variety of diseases and
XX disorders, depending on their function. Diseases that may be treated
XX include blood coagulation disorders, cancers and cellular adhesion
XX disorders. They may also be used to raise antibodies. AA233891 to
XX AA234330, and AA234685 to AA234774 represent polynucleotide and
XX polypeptide sequence given in the exemplification of the present
XX invention.
XX
XX Sequence 747 AA:

```

```

alignment_scores:
    Quality: 58.00      Length: 48
    Ratio: 2.762        Gaps: 1
Percent Similarity: 43.750 Percent Identity: 27.083

```

```

alignment_block:
US-09-712-768-5 x AA234758

```

```

Align seg 1/1 to: AA234758 from: 1 to: 747
7 CACGGCGACCGCAGAGAA..... 24
|||||
217 Htsglyaspargserlysgluserleuvalserphealactglnhsya 233
25 .....ACCGGATG 34
233 largserthvalthrghleutrpthrghlyasnphvalasnserileg 250
35 GCCTAGCGATTGCCATCGCGCTTGCGATCTTACCCGTGC 78
|||||
250 lntThrAlphealactgylcyltrpleuilethrphcys 264
seq_name: /SIDS2/gcgdata/geneseq/geneseq/AA2000.DAT:AA24314
seq_documentation_block:
ID AAB44314 standard; Protein: 747 AA.
XX
XX AAB44314;
AC
AC 08-FEB-2001 (first entry)
DT
DT Human PRO1012 (UNQ495) protein sequence SEQ ID NO:459.
DE
DE Human: secreted protein; transmembrane protein; PRO; EST; cytosolic;
XX expressed sequence tag; detection; cancer.
KW
KW Homo sapiens.
OS
OS Homo sapiens.
PN
PN WO200053756-A2.
XX
XX 14-SEP-2000.
PD
PD 18-FEB-2000; 2000WO-US04311.
PE
PE 08-MAR-1999; 99WO-US05028.
XX
XX 12-MAR-1999; 99US-0123957.
PR
PR 29-MAR-1999; 99US-0126773.
XX
XX 21-APR-1999; 99US-0130232.
PR
PR 28-APR-1999; 99US-0131445.
XX
XX 14-MAY-1999; 99US-0141037.
PR
PR 23-JUN-1999; 99US-0141037.
XX
XX 26-JUL-1999; 99US-0145698.
PR
PR 29-OCT-1999; 99US-0162506.
XX
XX 30-NOV-1999; 99WO-US28313.
PR
PR 02-DEC-1999; 99WO-US28351.
XX
XX 02-DEC-1999; 99WO-US28351.
PR
PR 16-DEC-1999; 99WO-US30095.
XX
XX 30-DEC-1999; 99WO-US31274.
PR
PR 05-JAN-2000; 2000WO-US00219.
XX
XX 06-JAN-2000; 2000WO-US00277.
PR
PR 06-JAN-2000; 2000WO-US00376.
XX
XX (GETH ) GENENTECH INC.
XX
XX Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;
XX Ferrara N, Filvaroff E, Fong S, Gao W, Gettier H, Gerritsen ME;
XX Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ;
XX Kljavin ID, Kuo SS, Napier MA, Pan J, Paoni NF, Roy MA;
XX Shelton DL, Stewart RA, Tumas D, Williams PM, Wood WI;
XX
XX WPI: 2000-611443/58.
XX N-PSDB; AAC78573.
XX
XX Novel PRO polypeptides and polynucleotides used in detection methods,
XX to target bioactive molecules to specific cells, and to modulate
XX cellular activities -
XX
XX Claim 12; Fig 190; 636pp; English.
XX
XX AAC78458 to AAC78599 represent polynucleotide and EST (expressed

```

CC sequence tag) sequences which encode secreted or transmembrane PRO
 CC polypeptides. The PRO polynucleotides and polypeptides have cytostatic
 CC activity. The polynucleotides and polypeptides can be used for detecting
 CC the presence of PRO polypeptides in samples, for linking bioactive
 CC molecules to cells and for modulating biological activities of cells,
 CC using the polypeptides for specific targeting. The polypeptide targeting
 CC can be used to kill the target cells, e.g. for the treatment of cancers.
 CC The polypeptide pairs provide specific targeting of bioactive molecules
 CC to cells. AAC78600 to AAC78987 represent PCR primers and probes used in
 CC the isolation of the PRO polynucleotide sequences.

XX Sequence 747 AA;

alignment_scores:

Quality: 58.00 Length: 48
 Ratio: 2.762 Gaps: 1
 Percent Similarity: 43.750 Percent Identity: 27.083

alignment_block:

US-09-712-768-5 x AAB44314 ..

Align seg 1/1 to: AAB44314 from: 1 to: 747

7 CACGCGACCGCAAGAA..... 24

217 HstglYAspArgSerLySGlSerLeuValSerPheAlaMetGlnHsVa 233

25ACCGGATTTG 34

233 LArgSerThrValThrgluLeuTrpHnglyAsnPhenValasnsertileg 250

35 GCCATGCGATTCGCTGCTTGGCTGGATCTTTACCCCTGTGC 78

250 lnrHAlaPheAlaIaGlylLeGlyTrpLeuIleTrpHncys 264

seq_name: /SIDS2/gcdata/geneseq/geneseq/AA2001.DAT:AAU12384

seq_documentation_block:

ID AAU12384 standard; Protein; 747 AA.

XX AAU12384;

XX 24-OCT-2001 (first entry)

DE Human PRO1012 polypeptide sequence.

XX Human secretory and transmembrane; PRO; mammalian; cancer; lung;

KW breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;

KW cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;

XX adipocyte; A-peptide; factor VIIA; gene therapy.

XX Homo sapiens.

XX WO200140466-A2.

XX 07-JUN-2001.

XX 01-DEC-2000; 2000WO-US32678.

XX 01-DEC-1999; 99WO-US28301.

XX 01-DEC-1999; 99WO-US28304.

XX 02-DEC-1999; 99WO-US28551.

XX 02-DEC-1999; 99WO-US28564.

XX 02-DEC-1999; 99WO-US28565.

XX 02-DEC-1999; 99WO-US30095.

XX 20-DEC-1999; 99WO-US30911.

XX 30-DEC-1999; 99WO-US30995.

XX 06-JAN-2000; 2000WO-US00277.

XX 11-FEB-2000; 2000WO-US03565.

PR 18-FEB-2000; 2000WO-US04341.
 PR 18-FEB-2000; 2000WO-US04342.
 PR 22-FEB-2000; 2000WO-US04414.
 PR 24-FEB-2000; 2000WO-US04914.
 PR 24-FEB-2000; 2000WO-US05004.
 PR 01-MAR-2000; 2000WO-US05601.
 PR 20-MAR-2000; 2000WO-US07377.
 PR 21-MAR-2000; 2000WO-US07532.
 PR 30-MAR-2000; 2000WO-US08439.
 PR 17-MAY-2000; 2000WO-US13705.
 PR 22-MAY-2000; 2000WO-US14042.
 PR 30-MAY-2000; 2000WO-US14941.
 PR 02-JUN-2000; 2000WO-US15264.
 PR 10-NOV-2000; 2000WO-US30873.

XX (GETH) GENENTECH INC.

XX Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;

XX Gerritsen ME, Goddard A, Godowski PJ, Gueney AL, Sherwood S;

XX Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WL, Zhang Z;

XX WPI: 2001-408281/43.

XX N-PSDB; AAS21456.

XX Isolated, secretory and transmembrane PRO polypeptide used to detect

XX other PRO polypeptides, link bioactive molecules to cells expressing

XX PRO polypeptides, and detect the presence of mammalian tumours e.g.

XX lung, breast, prostate, cervical

XX Claim 12; Fig 426; 813pp; English.

XX AAU12172-AAU12446 represent novel human secretory and transmembrane

XX PRO polypeptides. The PRO polypeptides are useful to detect other

XX PRO polypeptides, to link bioactive molecules to cells expressing

XX PRO polypeptides, to modulate biological activities of cells expressing

XX PRO polypeptides, and to detect the presence of mammalian lung, colon,

XX breast, prostate, rectal, cervical or liver tumours by comparing PRO

XX polypeptide expression in a cell sample to that in a control sample.

XX Some of the 275 sequences are also useful to stimulate the release of

XX tumour necrosis factor-alpha (TNF-alpha) from human blood, the

XX proliferation or differentiation of chondrocytes, the proliferation or

XX gene expression in pericyte cells, the release of proteoglycans from

XX cartilage, the proliferation of inner ear utricular supporting cells or

XX of T-lymphocytes, the release of a cytokine from peripheral blood

XX monocytes (PBMCs), or the proliferation of endothelial cells. Some of

XX the PRO polypeptides may modulate glucose or free fatty acid uptake by

XX skeletal muscle cells or by adipocytes, or inhibit binding of A-peptide

XX to factor VIIA. The PRO polypeptides can be used in assays to identify

XX molecules involved in binding interactions. The polynucleotides encoding

XX PRO polypeptides can be used to generate probes, antisense RNA/DNA,

XX transgenic or knock out animals and can be used in gene therapy.

XX Sequence 747 AA;

alignment_scores:

Quality: 58.00 Length: 48
 Ratio: 2.762 Gaps: 1
 Percent Similarity: 43.750 Percent Identity: 27.083

alignment_block:

US-09-712-768-5 x AAU12384 ..

Align seg 1/1 to: AAU12384 from: 1 to: 747

7 CACGCGACCGCAAGAA..... 24

217 HstglYAspArgSerLySGlSerLeuValSerPheAlaMetGlnHsVa 233

25ACCGGATTTG 34

233 LArgSerThrValThrgluLeuTrpHnglyAsnPhenValasnsertileg 250

35 GCCTAGCGATTGCATCGCCTTGCGTGCATCTTACCGTGTGC 78
 |||::||| |||::|||::|||::|||::|||
 250 InThraAlaPheAlaAlaGlyIleGlyTrpLeuIleThrPheCys 264

seq_name: /SID52/gcgdata/geneseq/geneseqp/AA2001.DAT: AAB92743

seq_documentation_block:

ID AAB92743 standard; Protein; 768 AA.

AC AAB92743;

DT 26-JUN-2001 (first entry)

DE Human protein sequence SEQ ID NO:11190.

KM Human; primer: detection; diagnosis; antisense therapy; gene therapy.

OS Homo sapiens.

PN EPI074617-A2.

PD 07-FEB-2001.

PF 28-JUL-2000; 2000EP-0116126.

PR 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

PA (HELI-) HELIX RES INST.

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

DR WPI; 2001-318749/34.

PT Primer sets for synthesizing polynucleotides, particularly the 5602

PT full-length cDNAs defined in the specification, and for the detection

PT and/or diagnosis of the abnormality of the proteins encoded by the

PT full-length cDNAs -

PS Claim 8; SEQ ID 11190; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602

CC full-length cDNAs defined in the specification. Where a primer set

CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary

CC to the complementary strand of a polynucleotide which comprises one of

CC the 5602 nucleotide sequences defined in the specification, where the

CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination

CC of an oligonucleotide comprising a sequence complementary to a

CC polynucleotide which comprises a 3'-end sequence, where the

CC oligonucleotide comprises at least 15 nucleotides; and the combination

CC of the present invention.

CC

CC

CC

CC

CC

CC

Ratio: 2.762
 Percent Similarity: 43.750 Percent Identity: 27.083
 alignment_block:
 US-09-712-768-5 x AAB92743 ..

Align seq 1/1 to: AAB92743 from: 1 to: 768

7 CACGCGCAGCCGAGAGAA..... 24

192 HisGlyAspArgSerIysGluSerLeuValSerPheAlaMetGlnHisIva 208

25ACCGCATG 34

208 IArgSerThrValThrGluLeuTrpThrGlyAsnPheValAsnSerIleG 225

35 GCCTAGCGATTGCATCGCCTTGCGTGCATCTTACCGTGTGC 78

225 InThraAlaPheAlaAlaGlyIleGlyTrpLeuIleThrPheCys 239

seq_name: /SID52/gcgdata/geneseq/geneseqp/AA2001.DAT: AAB93204

seq_documentation_block:

ID AAB93204 standard; Protein; 793 AA.

AC AAB93204;

DT 26-JUN-2001 (first entry)

DE Human protein sequence SEQ ID NO:12172.

KM Human; primer: detection; diagnosis; antisense therapy; gene therapy.

OS Homo sapiens.

PN EPI074617-A2.

PD 07-FEB-2001.

PF 28-JUL-2000; 2000EP-0116126.

PR 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

PA (HELI-) HELIX RES INST.

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

DR WPI; 2001-318749/34.

PT Primer sets for synthesizing polynucleotides, particularly the 5602

PT full-length cDNAs defined in the specification, and for the detection

PT and/or diagnosis of the abnormality of the proteins encoded by the

PT full-length cDNAs -

PS Claim 8; SEQ ID 12172; 2537pp + CD ROM; English.

XX The present invention describes primer sets for synthesizing 5602

CC full-length cDNAs defined in the specification. Where a primer set

CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary

CC to the complementary strand of a polynucleotide which comprises one of

CC the 5602 nucleotide sequences defined in the specification, where the

CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination

CC of an oligonucleotide comprising a sequence complementary to a

CC polynucleotide which comprises a 3'-end sequence, where the

CC oligonucleotide comprises at least 15 nucleotides and the combination

CC of the 5'-end sequence/3'-end sequence is selected from those defined in

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

alignment_scores: 58.00 length: 48

CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB9893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.

XX Sequence 793 AA;

alignment_scores:
Quality: 58.00 Length: 48
Ratio: 2.762 Gaps: 1
Percent Similarity: 43.750 Percent Identity: 27.083

alignment_block:

US-09-712-768-5 x AAB93204 ..

Align seg 1/1 to: AAB93204 from: 1 to: 793

7 CACGGCGACCGCAAGAAA..... 24
|||||||.....
217 HisGlyAspArgSerLysGluSerLeuValSerPheAlaMetGlnHisVa 233
25ACCGGATG 34
233 LArgSerThrValThrGluLeuThrPthrGlyAsnPheValAsnSerIleG 250

35 GCCTAGCGATTGGCATCGGCTGGCTGATCTTTACCGTGTGC 78
|||||||.....
250 InThrAlaPheAlaIaGlyIleGlyTyrPleuIleThrPheCys 264

seq_name: /SID2/gcdata/geneseq/geneseq/AAB95309

seq_documentation_block:

ID AAB95309 standard; Protein: 793 AA.

XX AAB95309;

XX 26-JUN-2001 (first entry)

XX Human protein sequence SEQ ID NO:17552.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX Homo sapiens.

XX EPI074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-0116126.

XX 29-JUL-1999; 99JP-0248036.

XX 27-AUG-1999; 99JP-0300253.

XX 11-JAN-2000; 2000JP-0118776.

XX 02-MAY-2000; 2000JP-0183767.

XX 09-JUN-2000; 2000JP-0241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayashi K, Yamamoto J,

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602
XX full-length cDNAs defined in the specification, and for the detection
XX and/or diagnosis of the abnormality of the proteins encoded by the

PT full-length cDNAs -
XX
XX Claim 8; SEQ ID 17552; 2537bp + CD ROM; English.

CC The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dr primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB9893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.

XX Sequence 793 AA;

alignment_scores:
Quality: 58.00 Length: 48
Ratio: 2.762 Gaps: 1
Percent Similarity: 43.750 Percent Identity: 27.083

alignment_block:

US-09-712-768-5 x AAB95309 ..

Align seg 1/1 to: AAB95309 from: 1 to: 793

7 CACGGCGACCGCAAGAAA..... 24
|||||||.....
217 HisGlyAspArgSerLysGluSerLeuValSerPheAlaMetGlnHisVa 233

25ACCGGATG 34

233 LArgSerThrValThrGluLeuThrPthrGlyAsnPheValAsnSerIleG 250

35 GCCTAGCGATTGGCATCGGCTGGCTGATCTTTACCGTGTGC 78
|||||||.....

250 InThrAlaPheAlaIaGlyIleGlyTyrPleuIleThrPheCys 264